

DOCUMENT RESUME

ED 160 598

95

SP 013 167

AUTHOR Suppes, Patrick, Ed.
 TITLE Impact of Research on Education: Some Case Studies. Summaries.
 INSTITUTION National Academy of Education, Washington, D.C.
 SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
 PUB DATE 78
 CONTRACT NIE-C-74-0116
 NOTE 119p.; For related document, see SP 013 168 ; Parts marginally legible due to small type
 AVAILABLE FROM National Academy of Education, 11 Dupont Circle, Washington, D.C. 20036 (Free)
 EDRS PRICE MF-\$0.83 Plus Postage. EC Not Available from EDRS.
 DESCRIPTORS *Abstracts; Behavior Theories; *Case Studies; *Educational Improvement; Educational Legislation; Educational Policy; Educational Psychology; *Educational Research; Government Role; Linguistics; Predictive Ability (Testing); *Research Utilization; School Architecture; Scientific Methodology; Vocabulary Development

ABSTRACT

Significant cases of research in terms of their effect on school practices have been extensively documented by educational researchers. This booklet summarizes nine such studies, which are published in detail in a companion volume ("Impact of Research on Education: Some Case Studies"). These studies have a variety reasonably characteristic of educational research. In presenting the various case studies, the authors have not focused on the scientific validity of the ideas they present but on the impact of these ideas on educational practice. Essays are written, individually and collaboratively, by the following authors: (1) John B. Carroll; (2) Geraldine Joncich Clifford; (3) Robert Glaser; (4) Guy J. Groen; (5) Hermine Warren; (6) Patrick Suppes; (7) H. H. Stern; (8) Marjorie Bingham Wesche; (9) Birgit Harley; (10) J. W. Getzels; (11) Torsten Husen; (12) Marvin Bressler; and (13) Joel S. Berke. (JD)

 * Reproductions supplied by EDES are the best that can be made *
 * from the original document. *

ED160598

SP

IMPACT OF RESEARCH ON EDUCATION: SOME CASE STUDIES

Summaries

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Edited by

PATRICK SUPPES

PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

NAE

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC) AND
USERS OF THE ERIC SYSTEM™

NATIONAL ACADEMY OF EDUCATION

Washington, D.C.

1978

SP013 167

237,7000

The project reported herein was performed pursuant to a contract (NIE-C-74-0116) from the National Institute of Education, U.S. Department of Health, Education, and Welfare. However, the opinions expressed do not necessarily reflect NIE position or policy, and no official endorsement by the National Institute of Education or the U.S. Department of Health, Education, and Welfare should be inferred.

See

Business correspondence should be addressed to
National Academy of Education
11 Dupont Circle
Washington, D.C. 20036

Copyright © 1978 by National Academy of Education
All rights reserved.

Printed in the United States of America

CONTENTS — SUMMARIES

PREFACE	v
<i>John B. Carroll</i>	
Chapter 1. ON THE THEORY-PRACTICE INTERFACE IN THE MEASUREMENT OF INTELLECTUAL ABILITIES	1
<i>Geraldine Joncich Clifford</i>	
Chapter 2. WORDS FOR SCHOOLS: THE APPLICATIONS IN EDUCATION OF THE VOCABULARY RESEARCHES OF EDWARD L. THORNDIKE	16
<i>Robert Glaser</i>	
Chapter 3. THE CONTRIBUTIONS OF B. F. SKINNER TO EDUCATION AND SOME COUNTERINFLUENCES	30
<i>Guy J. Groen</i>	
Chapter 4. THE THEORETICAL IDEAS OF PIAGET AND EDUCATIONAL PRACTICE	43
<i>Patrick Suppes and Hermine Warren</i>	
Chapter 5. PSYCHOANALYSIS AND AMERICAN ELEMENTARY EDUCATION	54
<i>H. H. Stern, Marjorie Bingham Wesche, and Birgit Harley</i>	
Chapter 6. THE IMPACT OF THE LANGUAGE SCIENCES ON SECOND-LANGUAGE EDUCATION	68
<i>J. W. Getzels</i>	
Chapter 7. PARADIGM AND PRACTICE: ON THE IMPACT OF BASIC RESEARCH IN EDUCATION	76

Torsten Husén

- Chapter 8.* EDUCATIONAL RESEARCH AND EDUCATIONAL REFORM: 84
A CASE STUDY OF SWEDEN

Marvin Bressler and Joel S. Berke

- Chapter 9.* EDUCATION, SOCIAL SCIENCE, AND THE LAW: 97
THE RODRIGUEZ CASE AND SCHOOL FINANCE REFORM

PREFACE

EVERY modern society has as a major concern the education of its citizens. In some developing countries the educational activities of the government consume as much as a third of the national budget. In the United States, such activities currently require more than a hundred billion dollars a year, although, as in the case of other forms of social and economic activity, only a slight effort is made to study the character of these educational activities and to understand them as intellectual, economic, or social processes. It is true that there has been a long tradition, even if a fragile one, of studying educational matters, but most members of the educational community will agree that educational research is a relatively minor activity in comparison to education as a whole.

All of us on occasion probably feel that there is little hope that research, given the small national effort devoted to it, will seriously affect practice. Such pessimism is not historically supported by the evidence, however. Research can have an impact in our society and it certainly does in many different ways. In the case of education, every large school system has as part of its central office staff some sort of research unit. The schools and colleges of education associated with institutions of higher education throughout the country are all charged with research responsibilities, and some of these are specifically written into the legislative charter of the institution. When the Office of Education was established by federal legislation more than a hundred years ago in 1867, the first section of the Act defined its chief purpose as one of "collecting such statistics and facts as shall show the condition and progress of education in the several states and territories, and of diffusing information respecting the organization and management of schools and school systems and methods of teaching."

In the hundred years since the founding of the Office of Education, a great deal of educational research has taken place, and much of it has been sponsored by the federal government. In a previous study entitled *Research for Tomorrow's Schools* (L. J. Cronbach and P. Suppes, Eds., New York, Macmillan, 1969) a committee on educational research of the National Academy of Education outlined some significant cases of research in terms of their effect on school practices. One example was the theory of mental tests and pupil classification and the ways in which the theory has come to be applied in American public schools; another concerned the influence of Peirce's

and Dewey's concepts of knowledge on the formulation of educational aims; a third was concerned with the contributions of Thorndike's theory of learning to the teaching of arithmetic in the early decades of this century.

A number of members of the National Academy of Education have expressed the view that it would be appropriate to embark on a more systematic presentation of the kinds of examples just cited. The present work summarizes nine such studies, which are published in detail in a companion volume of some seven hundred pages. These studies have a variety reasonably characteristic of educational research.

In presenting the various case studies the authors have not focused on the scientific validity of the ideas they present, but on the impact of these ideas on educational practice. Without doubt, searching questions about validity could be raised, as is always the case in science to one degree or another, but at a time when the relevance of basic research to practice is being challenged, or at least not strongly supported in many quarters, an examination of the contributions to education of earlier research seems timely.

In Chapter 1, John B. Carroll examines in some detail the development and present status of the scientific knowledge that lies back of the extensive use of psychological tests, especially tests of intellectual capacity. Carroll's detailed and objective examination of both theory and use of tests that measure intellectual abilities provides a desirable perspective on the many controversies that surround current testing practices.

In Chapter 2, Geraldine Joncich Clifford examines the impact of Edward L. Thorndike's vocabulary research on the actual content of written material used in schools. From a historical standpoint, Thorndike's work is a paradigm example of scientific research that has influenced detailed matters of curriculum. Clifford casts light on several factors that tend to promote or retard educational innovation as well as on the problems that arise in implementing research findings in school practice.

In Chapter 3, Robert Glaser examines the contributions of B. F. Skinner to education. That Skinner's research has had an impact on education in this country, especially in the decade and a half after World War II, is widely recognized, but the broad intellectual perspective that Glaser provides on Skinner's work and its implications, especially on more recent approaches to behavior modification in educational settings, will be found clarifying, even by those who consider themselves close to the subject.

In Chapter 4, Guy J. Groen examines the theoretical ideas of Piaget and their contributions to educational practice. Groen's lucid

account of Piagetian theory and its history of dissemination is almost unique in the literature on Piaget for its efforts to relate the theory to broader issues of scientific methodology. Groen's discussion of the implications for education of Piagetian theory should prove valuable to a variety of readers.

In Chapter 5, Hermine Warren and I examine the influence of psychoanalysis in American elementary education during the period 1900 to 1957. I believe that our chapter is the first detailed account of these matters and for this reason alone represents an interesting case study of the impact of one of the most important psychological theories of this century on American education.

In Chapter 6, H. H. Stern, Marjorie Bingham Wesche, and Birgit Harley analyze the impact of the language sciences on second-language education. They examine in considerable detail the developments in linguistics, psycholinguistics, and social linguistics over the past thirty years that have affected the teaching of second languages. They make a special effort to identify the impact of linguistic research on an extensive set of practical problems, ranging from the preparation of curriculum materials to the attitude of teachers.

In Chapter 7, J. W. Getzels examines a number of cases and a variety of issues that have been widely discussed on the relation of basic research to educational practice. One of his most interesting examples, quite different from others discussed in this volume, concerns the use of theory-oriented work in the architectural design of classrooms.

In Chapter 8, Torsten Husén analyzes in considerable historical detail the relation between educational research and educational reform in Sweden since 1945. He writes with the authority of one who has played a prominent role in these matters throughout the entire period.

In Chapter 9, Marvin Bressler and Joel S. Berke examine the relation between social science research and the law, especially as reflected in the *Rodriguez* case, and the issue of school-finance reform. As they emphasize, judicial resolution of educational issues has been a prominent feature of education in the United States in the last three decades, and the use of research results in the judicial decision-making process has been evident.

Although these nine studies by no means exhaust the range of instances in which research has influenced educational practice, they do provide a reasonable sample of some of the most prominent instances of impact. Educational research as a discipline scarcely existed prior to the 20th century. Some of the chapters are devoted to topics that are bound to be remembered historically as among the

most important examples of educational research in this century, but I like to think of the case studies presented here as representing only a sampling of the beginning of educational research in its early years of development. What the future holds should be brighter and better, because it will be able to build on the kind of work reported in this volume and, as that building takes place, the impact on practice should become more marked.

The framework for the work done on these nine case studies was provided by support from a contract (NIE-C-74-0116) from the National Institute of Education, U.S. Department of Health, Education, and Welfare. Arthur Melmed of the National Institute of Education has been helpful at all stages, from initial conception to completion of the printing of the present volume. Without his continued help and support, the present volume would not have come into existence.

This project was almost completed during my term of office as president of the National Academy of Education, which ended in October 1977. I cannot express too strongly my appreciation of the dedication of the staff of the Academy during my term in office, who helped extensively on this project. I mention especially William D. Hyde, Jr., who was executive secretary of the Academy from January 1975 to June 1977. Dianne Kanerva, who has had primary editorial responsibility for the volume, Pentti Kanerva, who has had main responsibility for the fast and efficient computer-based typesetting, and Bonnie Senko and Marguerite Shaw, who have both helped in innumerable ways in preparing the volume for publication.

Readers interested in the details of the studies as published in the volume mentioned earlier should write to the Washington office of the National Academy of Education, as shown on the back of the title page. The volume bears the same title as this summary.

Patrick Suppes

*Stanford, California
December 1977*

Chapter 1 — Summary

ON THE THEORY-PRACTICE INTERFACE IN THE MEASUREMENT OF INTELLECTUAL ABILITIES

JOHN B. CARROLL

*University of North Carolina
at Chapel Hill*

A PRIME goal of scientific discovery and technological development is the production of new knowledge—knowledge that can be disseminated through communications in scientific journals, monographs, books, and the like. This knowledge can be taken advantage of, generally, only at the cost of some special attention and study on the part of the user. Much of the output of psychological and educational research is of this character; that is, the fruits of this research are mainly writings that are intended to provide educators with guidance as to how various aspects of schooling should be conducted.

Often, however, particularly in the physical and biological sciences, scientific knowledge can become embodied in some kind of physical object or substance—a “long-lasting” razor blade, a color television set, or a new drug or vaccine. Most of the time, such products can be employed quite successfully by the ordinary layman who has not the faintest idea of the technology in metallurgy, electronics, biochemistry, or whatever, that went into their development and production.

At least some of the products of educational research are concrete physical things. One of these products—insofar as it may consist of a set of one or more physical objects like test stimulus arrays, response devices, or scoring keys—is the mental test. The mental test is a concrete product of the subdiscipline and technology known as psychometrics that is somewhat analogous to the knowledge-embodiment products of technology in the physical and biological sciences.

A mental test usually *looks* like a rather simple object, consisting, perhaps, of a series of tasks based on visual or auditorily presented verbal, mathematical, or symbolic materials. The scientific knowledge that may have gone into the selection and construction of these tasks, or that would be employed in the analysis of performances on these tasks, would hardly be evident from superficial examination of a mental test. No matter how simple (or apparently simpleminded) the test content may appear to be, the selection or devising of this content, and the analysis of people's performances on these tests, is undergirded with a substantial and respectable amount of knowledge and technology and, although the present theory and technology are highly sophisticated, there are reasonable grounds for questioning at least some of it and for thinking that new and profitable directions for research can be found.

The focus here is on tests of so-called intelligence, IQ, or mental ability—partly because of space limitations, but more because this is one of the most controversial fields of testing. Problems of reliability, validity, scaling, standardization, and the like have to be considered in connection not only with tests of mental abilities but also with tests of educational achievement, personality, interests, attitudes, motor skills, and the rest. But there are somewhat special considerations—linked to basic psychological issues about the nature of human capacities—that apply with particular force in the domain of tests of intellectual abilities.

Even within the field of mental ability testing, we must limit our scope considerably. Little consideration is given here to the clinical use of IQ tests to diagnose mental retardation and so-called learning disabilities. The emphasis is on "group" tests—that is, tests given to fairly large groups of people at a time and usually of the "paper-and-pencil" variety. Starting roughly in the 1920s, such tests have often been used in schools to yield measures of intellectual capacity to serve as a partial basis for ability grouping, selection for special classes, etc., and tests of this type have increasingly been used in connection with admission to higher education.

The history of the mental testing movement can be traced through two somewhat arbitrarily defined periods: (a) an "early" or "developmental" period beginning in the late 19th century, during which the foundations of theory and practice in psychometrics were laid down; and (b) a "modern" period starting around 1935 with the founding of the Psychometric Society, during which many refinements were made in the technology of testing and during which testing became a major enterprise.

THE PERIOD OF DEVELOPMENT

During the period that began with the publication in 1869 of Sir Francis Galton's early work *Hereditary Genius* and ended with the founding, around 1935, of the Psychometric Society and its journal *Psychometrika*, the science of mental testing progressed from its crudest beginnings to a stage of high development, leaving only matters of refinement for the period from 1935 to the present. It was during this earlier period that the major problems in the study of mental abilities were identified, the basic methodologies were developed, and tests of mental ability came into wide use.

Galton believed that mental ability could be revealed by performances on rather simple tests of the acuity of the senses, the speed of reaction, and the like. He can, in fact, be credited with the invention of the mental test, although it was only several years later (1890) that Cattell introduced the term *mental test*. From 1884 to 1890, Galton collected data from the general public on a variety of simple mental tests at the Anthropometric Laboratory that he established in London. Although it was about this time that he also developed the notion of correlation, Galton did not derive any important conclusions about individual differences from this work, except possibly the fact that these differences tended to be distributed according to the Gaussian law of error.

The American psychologist James McKeen Cattell became interested in individual differences in speeds of reaction, sensory discrimination, word association, and other simple mental tasks. Cattell set about applying such tests in his laboratory and postulated that these tests might be predictive of the scholastic success of college students. Unfortunately, Cattell's hopes were dashed when an analysis of his data showed that the scores of the tests that he had given several years earlier to Columbia University undergraduates had no practically useful correlations with the students' college grades. These findings were taken as persuasive evidence that mental tests of simple mental reactions had no promise as predictors of scholastic achievement nor, for that matter, as measures of anything like intelligence.

From the debates in the literature of the times, one would think that the mental-testing movement was being laid to rest. There were, however, activities elsewhere that presaged a very healthy development.

The German psychologist Ebbinghaus (1897) devised a "completion" test for measuring the effects of fatigue on children's school

performances. It required the child, using a guessing procedure that depended on comprehension of the total context, to fill in blanks representing deleted syllables or words in a prose passage. Ebbinghaus noted the substantial correlation of scores on this test with children's performance in school and with teachers' estimates of brightness.

The French psychologist Binet, however, may be credited with discovering an appropriate way to measure intelligence or mental ability. After a long period of investigating the simpler tasks studied by Galton, Cattell, and others, Binet came to the conclusion that tests comprising more "complex" tasks, resembling the mental activities of everyday life, were more promising. He developed "*une échelle métrique d'intelligence*" (Binet & Simon, 1905), a scale consisting of a series of tasks of increasing difficulty, each task representing the typical performance of children at a particular chronological age level. The tasks were highly varied, but most of them relied in some way on the understanding of language and the ability to reason with either verbal or nonverbal (spatial, numerical) materials. This scale became the basis for individual intelligence tests in a number of countries and languages.

Ebbinghaus had unwittingly demonstrated the feasibility of group tests of intelligence, and it was early recognized by a number of American psychologists that many of the kinds of mental tasks included in the Binet-Simon scale could be adapted for group testing. However, whereas the Binet scale called upon the examiner to find that point or region on the scale of task difficulty that separated generally successful performance of a child from generally unsuccessful performance, the group testing method involved simply asking the examinee to try as many "items" as he could and the score was computed as the number of items correctly answered. The resulting scores were regarded as being measured on a "point scale." Furthermore, group tests tended to call merely for recognition of a correct answer among several possible options. Exactly when the "objective" or "multiple-choice" form of item was first used is not clear, but this format quickly became characteristic of nearly all group tests of mental ability.

Group testing contrasted with Binet-type individual tests also in the fact that testing was done almost exclusively by means of paper-and-pencil printed tests. This often required that the examinee be able to read; it seems that test constructors blithely assumed that the examinees would all be equal in their ability to read and that only their ability to handle the material presented in printed form was at issue.

This, in fact, was apparently the assumption of the group of prominent American psychologists who were called together to devise ways of giving psychological examinations to Army recruits at America's entry into World War I. The Army Alpha Examination (in several alternate forms) was produced and administered to more than a million Army men, and subsequently Group Examination Beta was developed for illiterates and recruits who did not know English. Scores on both of these tests were scaled in terms of mental age, the concept originated by Binet. The concept of the IQ originated by Stern (1912) for assessing children's rate of mental growth was (correctly) regarded as inappropriate for use with this adult population.

The wide publicity given to group testing during World War I was attended with keen interest by educators and school administrators, who felt that group mental tests held great promise in helping teachers assess the learning capacities of their students. There ensued a flood of tests that imitated the Army tests in content and format. Notions of "standardization" (the derivation of norms for representative samples of the population for which a test was designed) and "validation" (establishment of substantial correlation of test scores with independent measures of the attributes being tested or predicted) were developed during this period. Beginning in 1925, Thurstone's *American Council on Education Psychological Examination for College Freshmen* yielded two scores, a "linguistic" score and a "quantitative" score; similar scores ("verbal" and "mathematical") were yielded by the Scholastic Aptitude Examinations constructed by C. C. Brigham for the College Entrance Examination Board (CEEB), starting in 1926. The culmination of this trend, for the period under consideration here, was represented by E. L. Thorndike's development in 1927 of his CAVD test, which contained four quite different parts (Completion, Arithmetic, Vocabulary, and Directions) and yielded scores on "altitude," "width," and "area."

These and many other tests became widely used in schools and colleges, and the publication and distribution of tests grew into a large, and often highly lucrative, enterprise. The concept of "intelligence" and "scholastic aptitude" came into general acceptance, but the notion that they reflected largely the effects of native endowment in interaction with schooling was, apparently, slow in coming.

THE MODERN PERIOD

The beginning of the modern period, as we are calling it, may with some justification be placed in the year 1935, the date of the founding of the Psychometric Society by a group of scholars who had been

called together by Louis Leon Thurstone and some of his colleagues at the University of Chicago. It was Thurstone and the leaders of the Psychometric Society, notably Paul Horst, who were primarily responsible for the origination and dissemination of the theoretical and technological developments that had a major impact on mental testing during this period.

By 1935, group mental testing had become widely adopted throughout all levels of education; by this time, also, it was beginning to be utilized in business and industry. Mental ability testing had become an established subject matter for research and teaching in nearly all the major universities, both in departments of psychology and in departments of education. Among the university centers of research in mental ability testing that were particularly prominent at the time were those at Columbia (especially at Teachers College), Stanford, Harvard, Minnesota, Ohio State, Iowa, California, and Chicago. (Centers of research in England and Scotland were also still very active.) Thousands, or more like millions, of "intelligence tests" were administered every year in elementary and secondary schools, often with a kind of blind faith in the scores. Increasingly, colleges and universities were using mental ability test results in their admissions and placement processes, although it was not until 1937 that the College Entrance Examination Board established scholastic aptitude tests as a regular part of its admissions testing program.

Most of the mental testing that was being done in 1935 was based on the view that "intelligence" is a more or less unitary trait. Most tests of intelligence yielded a single score for an individual—expressed usually in terms of an "intelligence quotient" or some derivative of the intelligence quotient. The common view was that intelligence test scores reflected some largely innate capacity of the individual to think in abstract terms, to learn in school, and to adapt to the requirements of an increasingly complex technological society. Intelligence tests were valued as scientific measurements to the extent that they showed high degrees of reliability and validity as indicated by the techniques developed by Spearman, Kelley, and other early scholars in mental testing.

Presaged in the later years of the testing movement's early period, but taking full form only in the 1940s and 1950s, two striking technological developments were initiated around 1935, chiefly by Thurstone and his colleagues at the University of Chicago: (a) multiple-factor analysis, a mathematical technique whereby it was clearly shown that "intelligence" is not a simple unitary trait; and (b) the refinement of the statistical theory of mental tests, leading to new ways of constructing mental tests and studying their measurement

characteristics. Both of these developments had a substantial impact upon the development and construction of mental tests during the modern period.

The history of testing throughout the modern period can be traced through the publications of a remarkable husband-and-wife team based at Rutgers University. It was shortly before 1935 that Oscar K. Buros, then an assistant professor who had been trained at Teachers College, became disturbed with the laxity of the standards by which mental tests were being published and distributed and decided to initiate a series of publications that would constitute a kind of "consumer's report" on mental and educational tests. The first two of these compilations were published by the School of Education at Rutgers, but by 1941 Buros and his wife Luella were able to establish their own independent nonprofit publishing house. Increasingly weighty tomes were issued in 1941, 1949, 1953, 1959, and 1972, with reviews of literally thousands of mental and educational tests; even at this writing, a final compilation is in preparation, the *Eighth Mental Measurements Yearbook*.

While the modern period saw the introduction of many innovations and refinements in the theory and practice of mental testing, these innovations and refinements were not utilized and applied as widely and as uniformly as they might have been. The Buroses may be given major credit for helping the field to police its own activities, but even the efforts of special committees of professional organizations such as the American Psychological Association (APA) and the American Educational Research Association (AERA) in publishing quasi-official technical standards for tests were not sufficient to ensure that all test-development activities conformed to these standards. Still less did the run-of-the-mill intelligence tests of the period take advantage of the technological refinements introduced during the period.

We can, however, discern the influence of Thurstone and other leaders in psychometrics in many published tests and in a variety of research and testing programs that appeared or occurred during the modern period. The prototype of the many "multifactor batteries" that were developed during the period was the *Primary Mental Abilities* (PMA) battery, which Thurstone based on the findings of his first major study of mental tests employing the multiple-factor analysis method (Thurstone, 1938). This study had identified at least seven somewhat independent "factors" of mental ability: perceptual ability, numerical ability, verbal ability, spatial-visualizing ability, memory, induction or generalizing ability, and deductive or reasoning ability. The importance of the PMA battery was that it first brought into general awareness the concept of differentiable mental ability factors.

Mental testing took new strides in the personnel programs of World War II. As had been the case in World War I, many well-known psychologists contributed their talents and energies to the speedy development and validation of tests suitable for use in the selection and training of millions of armed forces personnel for their varied tasks. These tests were fashioned along familiar lines and generally yielded single scores. Some of the armed services developed and validated tests to measure special aptitudes. Especially in the Army Air Force, Thurstone's ideas about primary mental abilities, and his methods of factor analysis, were utilized to develop batteries of special tests for the selection of pilots, navigators, bombardiers, and other classes of personnel. Also, in this and other testing programs, much was learned about improved and more efficient techniques of constructing tests, analyzing results, and using those results in personnel decisions.

After the war, J. P. Guilford, who had been in charge of much of the factor-analytic investigation done in the Army Air Forces, built upon this work to conduct, with the support of the Office of Naval Research, a 20-year "high level aptitudes" project to study the full range of mental abilities that might be measured (Guilford & Hoepfner, 1971). In the course of this work he conceived a new model of the structure of mental abilities, which he called "the Structure of Intellect" (SI; Guilford, 1956, 1967).

J. C. Flanagan, inspired by the promising results of factor-analytic investigations he had supervised in his wartime work, constructed a multifactor test battery called the *Flanagan Aptitude Classification Tests* (FACT). This battery was offered for general use through a commercial publisher, but Flanagan felt that a parallel battery should also be constructed and placed in the public domain so that long-range studies of the identification and utilization of human abilities could be conducted. Flanagan thus initiated the first major steps for the undertaking he called Project Talent, in which some 23 aptitude and achievement tests, as well as several questionnaires and interest inventories, were administered to about 140,000 high-school students around the country in March 1960 (Flanagan et al., 1962). There have already been several long-term follow-ups of the students in this sample, disclosing important new information about the growth of mental abilities and the way in which they interact with students' educational and occupational choices.

Commercial publishers were alert to the possibilities of developing multifactor batteries for wide use in secondary schools. One of the most highly regarded of such batteries is the so-called DAT, *Differential Aptitude Tests*, published by the Psychological Corporation. This

battery offers separate scores on Verbal Reasoning, Numerical Ability, Abstract Reasoning, Space Relations, Mechanical Reasoning, Clerical Speed and Accuracy, and Language Usage. Much labor was expended in item analysis, standardization, and validation of the tests, but little attempt was made to make the subtests "pure" measures of primary mental abilities; the goal was to develop tests that would have the greatest practical use in educational and vocational counseling in high schools.

Many other multifactor batteries were published during the period, and continuing interest in and demand for them is suggested by the fact that several entirely new batteries of this type have recently come to my attention, for example, the *International Primary Factors Battery* (Horn, 1973) and the *Comprehensive Abilities Battery* (Hakstian & Cattell, 1976), based on the continuing research in the identification of factors of mental ability.

In the meantime, there was movement toward the much greater use of mental ability tests in the college admissions process. The effects of the work of Carl C. Brigham in the experimental development of a scholastic aptitude examination were felt mainly in this modern period of mental testing. Brigham's early experiences included writing a doctoral dissertation on the Binet intelligence tests and service under Yerkes in the development of the Army psychological examinations in World War I. At Princeton in the 1920s, he developed a test that was used as early as 1925 in connection with admission to the university. Shortly thereafter, the College Entrance Examination Board turned to him to construct a similar examination for experimental use in a small number of "Ivy League" colleges. One of Brigham's innovations at that time was the use of "experimental sections" in the tests of successive years to explore the usefulness of new items and item types and to serve as a basis for making successive forms relatively equivalent in their derived scales. Brigham was evidently the first to develop systematic item-analysis procedures using large samples and also to emphasize the importance of "test security" (accounting for the return of all examination materials) in the administration of admissions tests.

Brigham carefully avoided calling his test an "intelligence test." Although he had earlier believed that tests like the Army Alpha measured intelligence and that the low average scores of certain immigrant populations indicated poor prospects for their success in America, he retracted these views in a famous article published in 1930 (Brigham, 1930). Brigham recognized the important difference between verbal and mathematical aptitudes. The Scholastic Aptitude Examinations (SAT), first adopted by the CEEB as a regular part of its

admissions examination program in 1937, have to this day included separate verbal and mathematical sections, each yielding a score on a scale that Brigham himself introduced—a scale with a mean of 500 and a standard deviation of 100 and with a range of 200 to 800, in a defined population of college admissions candidates.

Under Brigham's direction, the CEEB built a staff in Princeton to construct the successive editions of the SAT. This staff carried on after Brigham's death in 1943, continuing to meet Brigham's high standards of test construction. In 1947, it was merged into a new organization, the Educational Testing Service, which undertook the further development and administration of the CEEB examinations, including the SAT, as well as the construction and administration of many other testing programs. Educational Testing Service also became a major center of research on the theory of mental tests, with the addition of persons such as Harold Gulliksen, Ledyard Tucker, Frederick Lord, and William Angoff to its staff. Through the research of these specialists in mental test theory, the measurement characteristics of tests produced by ETS have been of particular excellence, although it cannot be said that *all* their research has yet been reflected in the actual development of tests.

Throughout the years of the modern period, the SAT gained wide acceptance. Whereas, for purposes of studying candidates' qualifications, the early Ivy League constituents of the CEEB were more interested in the results of essay achievement examinations in various subjects, these and other CEEB constituents increasingly found more usefulness in the SAT scores, which, many admissions officers concluded, were better indicators of the true promise of many capable students from unknown high schools who may not have been as well prepared (or coached) as their preparatory-school competitors.

The SAT did not continue to be alone in its field. In 1959, E. F. Lindquist of the University of Iowa, having resigned from his associations with the CEEB, established the American College Testing Program, which quickly grew to serve large numbers of colleges and universities, particularly in the Midwest and West. In contrast to the SAT, the ACT battery of examinations is much more heavily loaded with materials measuring generalized educational achievements in English usage, mathematics, social studies, and natural science. Above all, it seems to stress *reading* comprehension ability. According to a reviewer, "ACT suffers by comparison with the SAT in psychometric care and sophistication, is about equal in validity for predicting collegiate success, and excels somewhat in the variety and extent of ancillary services offered" (W. L. Wallace, in Buros, 1972, p. 615).

THE THEORY-PRACTICE INTERFACE IN PERSPECTIVE

The notion that individuals differ in intelligence, mental capacity, or mental ability, and that such differences can be measured and studied scientifically, has been one of the important ideas of Western civilization for more than a century. For better or for worse, it has been particularly influential in the shaping of educational policy and in the actual management and conduct of schooling. The study of the development of this notion affords an interesting case in the history of science. Yet, the complexity of the issues and the semantic overloads in the very term "intelligence" make it difficult to frame a discussion that is intended to be as judicious as possible on questions of what intelligence really is, what the relative influences of heredity and of environment on intellectual differences are, and how much mental abilities can be modified by human interventions. In this discussion, I have used the term "intelligence" in the broadest and most neutral sense.

When the concept of intelligence as an object of scientific study was introduced by Galton in the late 19th century, it was intimately associated with the notion that intelligence was in large measure an inherited trait. It was natural that Galton's studies of the breeding of plants and animals, and of the inheritance of certain physical characteristics in human beings, led him to believe in the hereditary nature of intelligence. It was probably this belief that caused him to neglect the role of learning and experience in the formation of mental ability.

After the first hesitant and largely unsuccessful attempts of Cattell and others to measure intellectual traits through simple sensory and perceptual tasks, the concept of intelligence started to have an influence on school practice with Binet's development of procedures for diagnosing the mental conditions and prospects of slow learners. While Binet did much to define intelligence as the manifestation of cognitive processes, he disavowed interest in the question of genetic origins of intelligence; in fact, he hoped that means could be found for improving the intelligence of mental defectives. Nevertheless, nearly all of those who used Binet's scale or its numerous adaptations were quick to assume that the mental growth rates that it measured were greatly influenced by heredity. Increased scientific respectability was lent to the concept of intelligence as an inherited trait by the statistical studies of the Galton-Pearson-Spearman school, and it is significant that the study of intelligence and its measurement was the principal catalyst for the further development of statistical methods

generally useful in the biological and behavioral sciences.

The gradual acceptance of a scientific concept of intelligence or mental capacity as a stable and probably inherited trait in individuals of all ages led to its use as a basis for mass testing, first in the selection of military personnel in World War I, and then in the schools. As an alternative to the individually administered Binet-type tests, group-administered paper-and-pencil tests were devised; in this process, subtle and largely unnoticed changes came about in the measurement properties and even the theoretical underpinnings of the tests. As methods for validating and standardizing tests became more refined, intelligence tests were eagerly seized upon by educators as a means of identifying promising students for selective schools and of assigning students to different levels of schoolwork difficulty. Intelligence testing was recognized as a scientific endeavor, and the notion of intelligence entered the public mind as a factor to be considered in the formulation of educational and social policies.

Examination and criticism of the concept of intelligence, either within or outside professional circles in mental testing, were rarely deep. Outside the profession, debate was often based on misunderstandings of the facts and theories offered by test developers, and in any case a general science of human characteristics and behavior had not progressed sufficiently to serve as a basis for sound public discussion of the implications of mental ability testing. Within the profession, argumentation centered around the generality and relative immutability of intellectual differences, but techniques and models for studying these matters were simplistic, and alternatives or supplements to genetic explanations of data were not thoroughly explored. Even in the public mind, it was hard to believe in environmental determinants of individual differences in intelligence in view of the obvious differences that could be observed among children in the same family, with presumably similar upbringing and advantages, and, at the same time, the obvious similarities in the mental capabilities of identical twins. Meanwhile, Brigham's recantation, in 1930, of his earlier hereditarian views and his development of a concept of "scholastic aptitude" distinct from intelligence were not well understood by either the mental testing profession or educators in general. Tests perceived to be intelligence tests—regardless of the names given them by their creators—continued to be produced in standard molds, periodically revised, and utilized in great quantities. Perforce, nearly all studies of the relevance of mental ability tests for selecting and assigning students in schools, or of the effects of genetic and environmental factors on test performances, used scores on global tests, of intelligence or scholastic aptitude as the variable of chief interest.

21

In the period since the middle 1930s, two lines of work have dominated the scientific study of individual differences: factor analysis and test theory. Factor analysis has been concerned with the decomposition of mental abilities into "factors," "primary mental abilities," or similar constructs. Test theory has been concerned with various statistical models and operations designed to produce certain desirable measurement properties in tests or test scores. In retrospect, these two types of scientific endeavor have had both productive and counterproductive features. Factor analysis has had at least the potentiality of facilitating the detailed understanding of what could be measured, and test theory has promoted the more accurate measurement of whatever could be measured. On the negative side, however, in factor analysis there has been too much interest in the identification of factors as end products of the research effort, and too little interest in explaining the factors identified in terms of their composition and antecedents. In test theory, there has been overemphasis on desirable measurement properties without corresponding consideration of what exactly was being measured—a pure trait or a mélange of correlated response tendencies. The two specialties drew apart, each failing to take adequate advantage of the achievements of the other. Both specialties developed in such technically complex directions that their findings and theories became all but unintelligible to the practical test user, and even to many persons honestly seeking to construct better and more valid tests.

Despite a few promising signs in recent research efforts, the present scene in intelligence testing is essentially one of stagnation, with much talk but little progress. Buffeted by public criticisms of testing and the shocks of controversy about alleged elitist and even "racist" motives of the testing movement, the profession has all but ceased working toward the development of better and more soundly based theories of intelligence and its uses. The overall decline in the utilization of intelligence tests in schools, the legal constraints on research with human subjects, the general public resistance to testing, for whatever reasons, and the inertia within the profession itself that has resulted from complacency about the sufficiency of presently available methodologies and testing procedures and disquiet about misuses of these procedures have made serious inroads on the ability of the testing profession to move forward. But large sectors of the profession are enmeshed in outmoded or at least questionable concepts of human behavior, particularly the *assumption* that mental abilities are relatively immutable even with extensive and prolonged experience or intervention and, furthermore, that they have almost overpowering genetic determinants.

It is my contention, nevertheless, that intelligence, if seen in a

proper light, is worthy of continued serious study, and that this study can have increasing relevance to problems of schooling and educational policy. The performances required on many types of mental ability tests have great and obvious resemblances to performances required in school learning, and indeed in many other fields of human activity. If these performances are seen as based on learned, developed abilities of a rather generalized character, it would frequently be useful to assess the extent to which an individual had acquired these abilities. This could be for the purpose of determining the extent to which these abilities would need to be improved to prepare the individual for further experiences or learning activities, or of determining what kinds and amounts of intervention might be required to effect such improvements. These determinations, however, would have to be based on more exact information than we now have concerning the effects of different types of learning experiences, including observation, practice, instruction, and so forth, on the improvement of these abilities. As matters stand now, we know very little about the parameters governing the growth of such individual attributes as language competence, reasoning ability, and speed of cognitive operations.

Even if performances on mental ability tests have at least some genetic determinants, as they probably do, they undoubtedly have substantial environmental determinants. Environmental variance is for the most part the only kind of variance that educators have a right to work with, but educators need to know the limits to which, and the conditions under which, this variance can be reduced. If the technological apparatus of factor analysis, test theory, and general experimental design can be directed toward the production of this kind of knowledge, and if this knowledge can be expressed and validated in a form that can be used by practitioners—a large order, perhaps, but feasible in principle—then the devoted efforts that have produced this apparatus will eventually have paid off.

REFERENCES

- Binet, A., & Simon, T. Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. *L'Année Psychologique*, 1905, **11**, 191-244.
- Brigham, C. C. Intelligence tests of immigrant groups. *Psychological Review*, 1930, **37**, 158-165.
- Buros, O. K. (Ed.). *The seventh mental measurements yearbook*. Highland Park, N.J.: Gryphon Press, 1972.
- Cattell, J. M. Mental tests and measurements. *Mind*, 1890, **15**, 373-381.
- Ebbinghaus, H. Über eine neue Methode zur Prüfung geistiger Fähigkeiten und ihre Anwendung bei Schulkindern. *Zeitschrift für angewandte Psychologie*, 1897, **13**, 401-459.

- Flanagan, J. C., Dailey, J. T., Shaycoft, M. F., Gorham, W. A., Orr, D. B., & Goldberg, I. *Design for a study of American youth, 1: The talents of American youth*. Boston: Houghton-Mifflin, 1962.
- Galton, F. *Hereditary genius: An inquiry into its laws and consequences*. New York: Appleton, 1869.
- Guilford, J. P. The structure of intellect. *Psychological Bulletin*, 1956, **53**, 267-293.
- Guilford, J. P. *The nature of human intelligence*. New York: McGraw-Hill, 1967.
- Guilford, J. P., & Hoepfner, R. *The analysis of intelligence*. New York: McGraw-Hill, 1971.
- Hakstian, A. R., & Cattell, R. B. *Comprehensive abilities battery*. Champaign, Ill.: Institute for Personality and Ability Testing, 1976.
- Horn, W. O. *International primary factors test battery*. Stevens Point, Wis.: International Tests, 1973.
- Stern, W. *Psychologische Methoden der Intelligenz-Prüfung*. Leipzig: Barth, 1912.
- Thurstone, L. L. Primary mental abilities. *Psychometric Monographs*, 1938, **1**.

Chapter 2 — Summary

WORDS FOR SCHOOLS: THE APPLICATIONS IN EDUCATION OF THE VOCABULARY RESEARCHES OF EDWARD L. THORNDIKE

GERALDINE JONCICH CLIFFORD

*University of California
at Berkeley*

FROM Edwin Doran—drawn to vocabulary research by the contention of a 19th-century grammar book that “a word is a sign of an idea” (1907, p. 431)—to Luella Cole—who wrote in a psychology textbook for prospective teachers that words are the “raw materials” of reading and thinking (1934, p. 47)—laymen and professionals have pondered the relationships of language to mental development.

The purpose of this essay is, first, to account for interest in a scientific study of words; second, to place in historical context the dominant position in vocabulary research of *The Teacher's Word Book* of Edward L. Thorndike; third, to identify the major findings and pedagogical implications of word-frequency research; and, fourth, to document the roles played by *The Teacher's Word Book* in emphasizing the readability of materials written for schools, while considering the differing characteristics of the several curriculum fields. The adequacy, elegance, rigor, or definitiveness of vocabulary and readability research is not the issue here. Rather, the plan is to describe and explain what happened in the word-world of elementary and secondary schooling in the period from 1920 to 1970 and to cast some light on the factors promoting or retarding educational innovation.

THE ORIGINS OF CONTEMPORARY INTEREST IN OBJECTIVE VOCABULARY DATA

Curious amateurs ordinarily preceded professionals in scientific investigating. In vocabulary studies, learned men counted the running words and different words in key pieces of literature and also estimated their own vocabularies. In 1907, Doran's bibliography listed 118 vocabulary studies and articles; by 1955, titles in vocabulary research numbered 2,601 (Dale & Reichert, 1957). In addition to psychologists and educators, other professionals—adult educators, librarians, government bureaucrats, mass communications specialists, military trainers—became interested in vocabulary research. In 1935, Gray and Leary used vocabulary calculations to estimate that one-half of available published reading matter was too difficult to be comprehended by half the adult population. In the 1950s, armed forces reading programs used readability formulas to select and grade books. Government, business, industry, and the press have tried to improve the readability of job applications, training materials, and other publications.

In an early study, Kirkpatrick (1891) estimated that adults of ordinary intelligence and common (elementary) school backgrounds averaged 10,000-word vocabularies; college graduates knew double that number. Present-day conclusions are that average vocabulary increases about 1,000 words per year between ages 3 and 20; those who thereafter increase their word knowledge much beyond learning proper names "can be called 'academic' persons" (Joos, 1964, p. 206).

Kirkpatrick's method was to sample systematically an abridged dictionary, test himself on the words selected, and project the proportion known to the proportion the list was of the dictionary's total number of words. In Thorndike's *The Teacher's Word Book* (1921a), vocabulary counters gained another source: words ranked by their frequency of occurrence in nearly 5 million running words of general reading. *The Teacher's Word Book of 30,000 Words* (Thorndike & Lorge, 1944) or random samples from an unabridged dictionary remain the chief bases for studies of vocabulary size.

Children's habits of oral language, knowledge of word meanings, and growth of reading and writing vocabularies were also being observed. Simple vocabulary counts predominated until the appearance of psycholinguistic research. The first important study of the recognition vocabularies of children used a sample of words drawn

from *The Teacher's Word Book* to suggest that the average 4-year-old child already knew about 1,500 words (Smith, 1926). Contemporary opinion is close to the early estimates. Chall (1967) writes, "The average first grader can probably use accurately and/or understand about four thousand different words" (p. 203).

Investigations of children's language sharpened the question: What does it mean to "know" a word? A crucial measurement issue was specifying how much meaning qualified as word knowledge. A function of *TTWB*¹ was to provide vocabulary test makers with a tool for selecting words of known utility by which the precision of knowledge of meaning of other words could be tested.

Thorndike, his early student B. R. Buckingham, and Henry Rinsland were prominent test makers as well as word counters. They personified the broader association of vocabulary research with the standardized testing movement which grew slowly between 1910 and 1920 and then exploded. When intelligence test makers found that a score made on their vocabulary subtests correlated most closely with the total score for the test, this provided an additional incentive for vocabulary research.

Tests furnished educators with more precise and sometimes sobering information about their students. They uncovered wide differences in pupil achievement—differences of from 4 to 8 years in a single classroom. Research clarified the role that "word power" played in reading achievement, a useful undertaking since reading disability was the widely acknowledged major cause of personal pupil failure and of shocking inefficiency in the system. Hilliard (1924) found that intelligence ranked first in importance in reading comprehension, and extent of meaning vocabulary second; giving training in vocabulary, he declared, is "well within the grasp of every teacher" (p. 40).

Between 1910 and 1930, two fundamental changes were effected in orthodox opinion on reading instruction. One shifted the focus from oral to silent reading. The second was a new emphasis on comprehension, which eclipsed other objectives of reading. The conception of meaningful reading mandated that attention be paid to children's background of experiences. The most meaningful reading will contain words and word relationships already part of the student's recognition (hearing) and oral (speech) vocabularies. By the 1930s educators were demanding of publishers that they produce schoolbooks of simpler vocabulary. Already, many publishers

¹ *The Teacher's Word Book* will hereafter be referred to as *TTWB*. There was a 1921 edition, a 1931 edition, and the 1944 Thorndike-Lorge edition.

explained the vocabulary "burden" of their textbooks as being "based on the Thorndike List" (Durrell & Sullivan, 1938, p. 192).

Thorndike began his own counts of words found in all sorts of reading materials around 1911, probably aware of the international flurry of activity, around 1900, in identifying essential words for the use of shorthand writers and foreign-language teachers. Dissatisfaction with spelling achievement and the unpopularity of spelling with teachers and students also generated lists, like Ayres's (1913b) and Nicholson's (1914); Ayres proposed using word frequency as the criterion in establishing a core of words for instruction but thought his own list insufficiently reliable.

STUDIES IN WORD USAGE: THE WORD COUNTERS

The singular advantage of *The Teacher's Word Book* in advancing subsequent vocabulary research was its criterion of word selection. Earlier word lists used "word difficulty" or the "technical importance" of words—both hard to define—as their principle of selection. Thorndike used frequency of occurrence in ordinary reading material, causing his work to endure as a research tool in psychology, education, and communications. A count of references to Thorndike's word studies, found in the 1973 and 1974 volumes of the computer-based *Social Sciences Citation Index*, totaled 234; the combined count of citations of other prominent compilers of word lists and architects of readability formulas was 53.

Another measure of the influence of *TTWB* is the impetus it gave others to create better frequency lists or those tapping other sources and patterns of word usage. Over 100 frequency-count word lists were subsequently published. The majority were local or otherwise transitory, notable primarily in raising awareness of vocabulary issues in the many school communities where list makers were using teacher and student help. A few particularized lists were important, used by educators, textbook authors, and Thorndike himself in constructing later editions of his *Word Book*. Not, however, until computers were available to handle many millions of words did new general lists appear.

Experienced educators quickly noted that words on the same level of frequency in *TTWB* might pose different degrees of conceptual difficulty for children and adults. Perhaps because of primary-grade teachers' greater receptivity to vocabulary issues and the continuing disagreements about methods of teaching beginning reading, the most significant early count of children's word usage was sponsored

by the International Kindergarten Union (M. D. Horn, 1928). Analyzing 6 million words of children's writing, Rinsland published *A Basic Vocabulary of Elementary School Children* (1945), with each word's frequency shown for each grade level.

That word difficulty is more than a matter of frequency became apparent when researchers examined "technical words"—words employed in curriculum fields but not found in *TTWB*. Even in school subjects not thought to possess a large technical vocabulary, word counters—armed with *TTWB* or the proliferating supplementary lists—reported an overdose of unfamiliar language of conceptual difficulty. Because technical words appeared so infrequently in general reading and were given insufficient repetition in textbooks, teachers were continually advised to provide such repetition and opportunity for practice.

CONCLUSIONS DRAWN FROM FREQUENCY RESEARCH AND THEIR IMPLICATIONS FOR EDUCATION

The following conclusions about word frequency emerged as a consensus of informed opinion within the two decades after 1920. Furthermore, school practices, especially in the language arts, came to show some consistency with the implications of these findings.

An early goal of vocabulary research was to establish the relationships between the vocabularies of hearing, speech, reading, and writing. The conclusion was that, as children mature and are educated, their several vocabularies converge; this is especially true of the words of highest frequency. Therefore, words helpful in one language activity in the classroom have a high probability of functioning well and of being meaningful in another. This supported a breaking down of compartmentalization in the curriculum. The unified "language arts" approach grew to popularity in the 1940s and 1950s and became the orthodox view in elementary schools.

As frequency counts multiplied, comparisons of their findings tested the presumption that children possess their own vocabulary. A significant overlap was found between the commonplace writing words of children and adults; the least agreement between children's and adults' vocabularies is in the nouns they use. Thorndike (1921b) early suggested that "by elimination of certain specially childish or 'literary' words from the first 500 of our list and the addition of certain words of special importance to the newcomer to America . . . we shall have a basic list of great value in teaching foreign adults to read English" (p. 364). An international community of English-language teachers used *TTWB* to prepare limited vocabulary

textbooks and apply readability formulas to grade English-language materials. In an era of worldwide immigration and expanding colonial empires, when the radio and phonograph crossed political borders with language, an interest also arose in finding or creating an "international language." Esperanto was one candidate; another was "Basic English," an 850-word vocabulary.

Vocabulary research confounded those who expected that word frequencies would distribute themselves along a "normal curve": a few words seldom used, a great number of middle-frequency words clustered together, and a few words regularly used. Instead, a drastically skewed distribution was found: a very few, very high frequency words; a small number of medium-frequency words; and a very large number of very low frequency words (Twaddle, 1966). The 300 most common words constituted 75% and the 1,000 commonest constituted 90% of the words we write. Since most textbooks contained a great many different words, they were tapping many of the less common, less important words. More seriously, they made inadequate provision for repetition of these less common words—and learning suffered. Hence, the first priority in teaching should be teaching wide-range words—"range" referring to the coverage a word receives over a number of sources. A manifestation of the vocabularists' influence was the growth of a common textbook vocabulary within and across the various school subjects. Second, educators were advised to reduce the "vocabulary burden" of almost all textbooks. Where older spelling books might contain 15,000 words, the total spelling curriculum of the elementary school shrank to 3,000 or 4,000.

The 10 most common words in adult correspondence in American English around 1910 were *the, and, of, to, I, a, in, that, you, for* (Ayres, 1913a). The 10 most common words in a sampling of American English printed in 1961 were virtually the same (Francis, 1967). These high-frequency words are grammatical words (also called service or function words), as distinguished from lexical words. Because they make up the structure of modern English prose, they are among the oldest words in the language. The Kučera-Francis (1967) computation found the most frequent 22 words were all grammatical words; nearly one of every three words you will read is one of these 22 words. The first lexical word in frequency was in the 64th position; it was the word *new*. The word *elephant* (a lexical word) is easily learned; *with* (a grammatical word) is not. While gathering statistical evidence of these words' frequencies did not invent the need to teach them, it underscored that need, with a probable increase in planned practice of the harder-to-learn words.

Another and related conclusion of vocabulary research was that

differences in vocabulary due to gender, geography, intelligence levels, and other factors do not much influence the base vocabulary of the 3,000 or 4,000 most common words (Greene, 1941, 1950). The basal readers prepared since 1960 may feature an ethnically integrated, less "middle-class" and suburban cast of characters, but they still conform to the orthodox view of a controlled vocabulary of common words, emphasizing meaningfulness and repetition. The restriction, through language, of school content more clearly to a culturally undifferentiated core was, after all, appropriate to a mass and universalizing school system.

From 1930 to 1937, 60% of the foreign-language investigations were vocabulary studies. Language specialists participated in using frequency data to evaluate school materials, to grade books, and to secure simplified textbooks. In 1939, Ernest Horn confirmed the striking agreement between the relative frequency of English words and their foreign-language equivalents; *time* is the most common noun in English and *Zeit* in German (E. Horn, 1939). Comparable frequency findings tended to support claims that word-frequency data possess validity and reliability.

Data acquired for the Lorge-Thorndike Semantic Count (Lorge, 1937) revealed a positive correlation between a word's frequency and its number of recognized (dictionary) meanings. Some of the extremely common words of multiple meanings (Bachmann, 1944) were *made* (25 meanings), *broken* (15), *fall* (14), and *touch* (11). Thus, the challenges to teachers lay both with the uncommon words' rarity and the common words' richness of meanings.

ON HOW AND WHY THINGS CHANGED IN SCHOOLS

By change is meant, first, the altering of school reading matter to improve its readability and, second, the concomitant heightening of acceptance of the view that learning words is a sometimes difficult task whose accomplishment cannot be left to chance. Word awareness may be manifested by augmenting the textbooks' word treatment, using workbooks for word-building purposes, and testing students to verify their word knowledge. The purpose here is to draw together generalizations about the forms through which Thorndike made his primary contributions and to advance some propositions about the factors that, together, defined the scope of change. That the vocabularists' successes were neither complete nor uniform across grade levels and curriculum areas suggests that conditions favoring receptivity, resistance, or disinterest merit analysis. While the historical constellation of variables and forces was specific as to time and place and

identity of the participants, some general enlightenment of the processes of institutional change may emerge from this analysis.

The Media of Thorndike's Influence: Tests and Textbooks

As the making of tests created new reading material after 1920, it was predictable that the selection and arrangement of test words would be subjected to scrutiny similar to that accorded textbook words. Despite the scarcity of published research on the subject, substantial evidence exists that test authors and publishers considered and used frequency word lists and, later, readability formulas. Test reviews and test bibliographies demonstrate Thorndike's impact upon test vocabulary. *TTWB*, his dictionaries, or other "Thorndike words" were employed in test construction—especially in reading, spelling, arithmetic, clerical, and vocabulary tests. Word research was used to select the words of test content and instructions, to choose words to act as distractors, and to arrange items in order of ascending difficulty.

Several books published for schools and teachers on their own construction and evaluation of tests spoke of appropriate language and careful attention to vocabulary in the test directions and content (Kelley & Krey, 1934; Bolton, 1943). *TTWB* was used to discredit New York State's system of testing the literacy of new voters, and a standardized test was constructed to replace the method of reading aloud extracts from the state constitution.

Many Americans do not realize that textbook use is more prevalent in this than in other systems. Three *Yearbooks* of the National Society for the Study of Education, between 1921 and 1931, challenged the suitability of textbooks. With *TTWB* came a new criterion, but one built on this culture's valued principle of "social utility": How does textbook language compare with the general language, using word frequency as the measure? Given the quest for precision that marked the era, Thorndike's impact was considerable.

Despite other pressures—the textbook adoption system, censorship, cost—competition led publishers to claim that the design of their textbooks was compatible with scientific vocabulary research. Textbook authors came increasingly from university departments of education and from educational research. Prominent vocabularists or their students authored most of the major publishers' series, especially in reading, spelling, and elementary-school mathematics—viz., Horn, Gray, Buckingham, Dolch, Yoakam, Betts, Powers, N. B. Smith, Gates, and Hildreth. The secrecy of techniques used in appraising and rewriting manuscripts led Chall (1958) to call readability measurement almost a "secret tool" of publishers.

An indisputable fact about 20th-century textbooks is the great amount of reading matter they contain. Charles Eliot calculated in 1898 that the entire 6-year elementary-school curriculum could be read in 46 hours; the brief textbooks then used were modeled after the catechism and intended to be memorized, not read. By 1936, however, Charles Judd reported a tenfold increase in reading volume. Given the requirement of adequate repetition of vocabulary and the greater knowledge available about technical vocabulary, textbooks in the content subjects also increased their required reading. Even the harshest critics of modern high-school English anthologies commended "their quantitative sufficiency of reading matter" and generally adequate vocabulary study (Lynch & Evans, 1963, p. 409).

Accompanying sheer increases in textbook bulk came additional exposure to printed words through the appearance of supplementary books: associated readers, enrichment readers, simplified readers, "adventure" books. Given this flood of material, one could anticipate a recurrence of the fluctuation in vocabulary practices characterizing earlier textbooks. That "great and unwarranted disparity among the vocabularies of books for children" was prevented was credited to the influence of *TTWB* (McKee, 1937).

Graded and vocabulary-controlled series of basal textbooks emerged in science, social studies, and arithmetic, allowing for vocabulary coordination across the curriculum. The workbooks that appeared and proliferated after 1926 provided for more systematic introduction and repetition of new and uncommon vocabulary. The teacher's manuals, also grown in size and specificity, further identified vocabulary loads, probable word difficulties, and methods to enhance comprehension. Considering how much teachers' opinions were formed by their manuals, this attention to vocabulary was crucial.

What Makes Research Influence Practice?

By the 1930s vocabulary study was so common that researchers spoke imperatively of the urgency of carefully controlling school vocabulary. Yet, judging from observed variability across fields, the effects of such exhortation were mixed. In this section of the chapter an attempt is made to explain the differential effects of vocabulary research and to understand the qualities of a research tradition that enhance or limit acceptance of educational research.

A "Logical Fit"

It is hypothesized that curriculum areas that are obvious "word fields" would be more responsive to vocabulary research. Spelling, reading, typing, shorthand, and foreign-language teaching (at least in its initial stages of instruction) are word fields. Each accepted the

principles of low rates of vocabulary addition and frequent repetition of old and new words. Within a quarter century, the vocabulary burden of basal readers was halved, and greater word repetition was provided. Awareness of high mortality rates in foreign-language study intensified the sense that little time existed to attempt the nonessential. The discovery of the skewed frequency distribution gave rise to an important pedagogical practice for shorthand teachers: The teacher emphasizes memorization of the first few hundred most frequent words; thereafter students must rely upon acquired skill in constructing, on the spot, legible outlines for those seldom-encountered words which run into the many thousands. Various frequency lists, used for research into teaching methods and in constructing connected materials for drill, also characterized typing courses; by 1961, over half of all ninth graders took a typing course.

In the foregoing subjects, teachers were basically in sympathy with vocabulary researchers' efforts to identify an essential word list of limited scope and to clarify the processes and parameters of vocabulary development. In contrast, there exist teaching fields, such as the fine arts, where teachers and authors of texts and courses of study have been suspicious of such educational research (see, e.g., Clifford, 1973).

Opposition stemmed partly from the utilitarianism associated with research. In teaching literature, for example, practitioners denied the "logic" of vocabulary research, although reading is the essence of the classroom experience. Evans and Lynch (1960) drew a distinction between reading material and literature. They were annoyed to find that 64 words in an extract from *Great Expectations*, some of which "rank as sixth-grade words on the Thorndike list, are sometimes replaced by what Thorndike classed as third-grade words, although the volume is intended for the ninth-grade" (Lynch & Evans, 1963, p. 51). While readability research apparently made few converts among literature teachers, these authors thought it had prospered lamentably well in the textbook publishing houses.

Status Factors

An omnibus category of explanation for the differential effects of vocabulary research has to do with relative status or security of place in the social structure of the schools. The theory offered is that a level of school or curriculum area will be differentially receptive to, or ready for, educational innovation on the basis of sociological dynamics generated by status considerations. A few examples should clarify the several dimensions of status.

In general, it has been and remains more prestigious to teach certain subpopulations—older students rather than younger, boys

than girls, the academically quick than the slow, advanced than beginning students, and selected pupils rather than all youngsters. It is in these prestigious programs that greater resistance to change is expected; teachers of these programs have privileges to protect and lesser reason to court the uncertainties accompanying change. In basic courses—required, preparatory, and frequently unpopular—teachers have less to lose and may seek a substitute identity, perhaps the reputation of being creative and forward-looking. It is well known that the subjects and teachers of elementary schooling were earlier and more accepting converts to progressive education than was true in secondary schools.

A second type of status issue, "status anxiety," was manifested by school geography, which in the 20th century was so beset by a succession of distracting "identity problems" that one marvels that geography teachers paid any attention to vocabulary research. Like other subjects, geography was urged to simplify its textbooks, but a greater concern was the threat to its very existence under the pressure to merge geography into the "social studies"; this threat commanded most of geography's attention through the 1930s.

Algebra illustrates a third kind of status consideration conditioning the movement of research into schools. In this case, it is algebra teachers' self-perceived status as mathematics teachers that apparently limited their receptivity to educational research. Algebra teachers value their status as experts in mathematics over any commitment to a learner-centered psychology. Conservatism in the teaching of algebra is also encouraged by the isolation of algebra from the pressures of mass education. Except for the college-bound high-school population—until recently a small minority of even high-school graduates—prospects for failure in traditional algebra can be counseled into general or business mathematics. Thus, algebra teachers have been denied considerable firsthand experience with the question of the contribution of language to the difficulties of their subject matter.

A Cumulative Research Tradition

Reflecting upon whether vocabulary and readability research displayed the cumulative properties that are believed to enhance the potential of research to affect everyday behavior, it appears that it passed this test rather well. There is an impressive continuity from the earliest, least mathematical word counts to complex and linguistically aware readability formulas. Furthermore, that the existing research corpus has generated rather precisely defined areas of still-needed measurements of other dimensions of readability should be acknowledged; this is one of the dividends of research cumulation.

The tradition of readability research represents an intricate linkage or network of extensions and modifications of key studies. In graphic form this network would resemble more a family tree of research than a pile of studies. In large measure, behind this conceptual and empirical network lies a personal-historical web: a system of communication of interests between teachers and their students. Any attempt to represent this web of personal associations in vocabulary and readability research shows the massive centrality and insistent outreach of Thorndike and Teachers College, Ernest Horn and the State University of Iowa, and W. S. Gray in reading and Guy Buswell in arithmetic vocabulary at the University of Chicago. Thorndike's seminal position increases when one recalls that Horn and Gray came under Thorndike's influence at Teachers College. Even an ostensible maverick, Rudolph Flesch, studied at Teachers College in the Readability Laboratory in the 1940s.

Social Sanction

As social agencies, public schools must be answerable, at least in part, to basic cultural imperatives and to various social pressures. By the 20th century it was widely agreed that public education represented a vital social investment, that popular aspirations and collective well-being demanded that it also be universal. Educational efforts aimed at schooling "all of the children of all of the people" had wider support than those that hinted of elitism or exclusion.

The schools by 1920 were receiving and retaining larger proportions of the less "word-facile" of the population. The economics of mass textbook publishing and school district purchasing promoted books aimed at the new "average reader"—the result being a simpler textbook than was formerly the standard. Moreover, research even showed that better readers also profited—in retention, reading rate, and attitude—when readability was improved.

Teachers reportedly preferred materials that increased student interest and lessened the need for explanation—books and courses of study that better fit the range of their students' abilities and backgrounds. A prominent reading and vocabulary expert, who made explicit the relationship of readability to a more egalitarian education, was Yoakam (1954). He characterized those who espoused school-books written to meet the canons of scholarship as having "an aristocratic rather than a democratic theory of education" (p. 430).

Timeliness

A function of educational research too little considered and appreciated is that of legitimization: giving the support of "facts" and the

prestige of science to existing trends and movements. Vocabulary research nicely complemented the teaching and curriculum reforms advocated by the visionaries of the early progressive education movement. Thorndike's debunking of the useless and recondite in traditional textbooks was more than an attack upon uncommon and unimportant words. It was part of a larger campaign against a false psychology of mental discipline, against the scholar's arrogance and the teacher's bad habits, and on behalf of clearer specifications of the social utilities of schooling. Thus, supporters of the "new education" received welcome help from the "vocabularyists" in their own struggles to "vitalize" the curriculum, emphasize the meaningful in classroom experiences, recognize and provide for group and individual differences, and equip students for a socially conscious adulthood.

REFERENCES

- Ayres, L. P. *A measuring scale for ability in spelling*. New York: Russell-Sage Foundation, 1913. (a)
- Ayres, L. P. *The spelling vocabularies of personal and business letters*. New York: Russell-Sage Foundation, 1913. (b)
- Bachmann, H. M. A semantic study of the books of two authors dealing with classical antiquity. *University of Pittsburgh Bulletin*, 1944, **40**, 16-27.
- Bolton, F. B. A study of vocabulary in the social studies. *Social Education*, 1943, **7**, 17-18.
- Chall, J. S. *Readability: An appraisal of research and application*. Columbus: Ohio State University, 1958.
- Chall, J. S. *Learning to read: The great debate*. New York: McGraw-Hill, 1967.
- Clifford, G. J. A history of the impact of research on teaching. In R. M. W. Travers (Ed.), *Second handbook of research on teaching*. Chicago: Rand-McNally, 1973.
- Cole, L. *Psychology of the elementary school subjects*. New York: Farrar & Rinehart, 1934.
- Dale, E., & Reichert, D. *Bibliography of vocabulary studies* (Rev. ed.). Columbus: Ohio State University, 1957.
- Doran, E. W. A study of vocabularies. *Pedagogical Seminary*, 1907, **14**, 401-438.
- Durrell, D., & Sullivan, H. B. Vocabulary instruction in the intermediate grades. *Elementary English Review*, 1938, **15**, 138-145, 160, 185-198.
- Evans, B., & Lynch, J. J. *Dialogues on the teaching of literature*. New York: Bookman, 1960.
- Francis, W. N. The Brown University standard corpus of English: Some implications for TESOL. In B. W. Robinett (Ed.), *On teaching English to speakers of other languages*. Washington, D.C.: Teachers of English to Speakers of Other Languages, 1967.
- Gray, W. S., & Leary, B. E. *What makes a book readable?* Chicago: University of Chicago Press, 1935.

- Greene, H. A. English—Language, grammar, and composition. In W. S. Monroe (Ed.), *Encyclopedia of educational research*. New York: Macmillan, 1941.
- Greene, H. A. English—Language, grammar, and composition. In W. S. Monroe (Ed.), *Encyclopedia of educational research* (Rev. ed.). New York: Macmillan, 1950.
- Hilliard, G. H. Probable types of difficulties underlying low scores in comprehension tests. *University of Iowa Studies in Education*, 1924, 2(6).
- Horn, E. The validity and reliability of adult vocabulary lists. *Elementary English Review*, 1939, 16, 129-134, 138.
- Horn, M. D. *A study of the vocabulary of children before entering the first grade*. Baltimore: International Kindergarten Union, 1928.
- Joos, M. Language and the school child. *Harvard Educational Review*, 1964, 34, 203-210.
- Judd, C. H. The significance for textbook-making of the newer concepts in education. *Elementary School Journal*, 1936, 36, 575-582.
- Kelley, T. L., & Krey, A. C. *Tests and measurements in the social sciences* (Report of the Commission on the Social Studies, Part 4). New York: Scribner, 1934.
- Kirkpatrick, E. A. Number of words in an ordinary vocabulary. *Science*, 1891, 18, 107-108.
- Kučera, H., & Francis, W. N. *Computational analysis of present-day American English*. Providence, R.I.: Brown University Press, 1967.
- Lorge, I. The English semantic count. *Teachers College Record*, 1937, 39, 65-77.
- Lynch, J. J., & Evans, B. *High school English textbooks: A critical examination*. Boston: Little, Brown, 1963.
- McKee, P. Word lists and vocabulary difficulty in reading matter. *Elementary English Review*, 1937, 14, 241-245.
- Nicholson, A. *A speller for the use of the teachers of California*. Sacramento, Calif.: State Printing Office, 1914.
- Rinsland, H. D. *A basic vocabulary of elementary school children*. New York: Macmillan, 1945.
- Smith, M. E. An investigation of the development of the sentence and the extent of vocabulary in young children. *University of Iowa Studies in Child Welfare*, 1926, 3(5), 1-93.
- Thorndike, E. L. *The teacher's word book*. New York: Teachers College, 1921. (a)
- Thorndike, E. L. Word knowledge in the elementary school. *Teachers College Record*, 1921, 22, 334-370. (b)
- Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words*. New York: Teachers College Press, 1944.
- Twaddle, W. F. Linguists and language teachers. In C. J. Kreidler (Ed.), *On teaching English to speakers of other languages*. Champaign, Ill.: National Council of Teachers of English, 1966.
- Yoakam, G. A. Unsolved problems in reading. *Elementary English*, 1954, 31, 427-430.

Chapter 3 — Summary

THE CONTRIBUTIONS OF B. F. SKINNER TO EDUCATION AND SOME COUNTERINFLUENCES

ROBERT GLASER

University of Pittsburgh

EDUCATORS in America, especially over the last 75 years, have searched, in one way or another, for psychological principles of human behavior from which educational practice could be derived. At the beginning of this century, education and psychology were closely linked in the work of many psychologists. Individuals like Edward Thorndike and John Dewey epitomized the close working relationship between the science of psychology and educational practice. After the time of Dewey and Thorndike, however, education and psychology went their separate ways. Psychology moved into the laboratory and focused on experimental work and theoretical development. Education, on the other hand, found its challenges with practitioners—with teachers and the professional educational establishment.

B. F. Skinner moved into this scene at a time when a strong relationship between psychology and education was beginning to be renewed. This chapter describes the part Skinner has played in the relationship between psychology and education and the effect of his ideas on educational practice. The chapter begins with a description of Skinner's theories of behavior and then describes how these theories have been applied in the development of teaching machines, programmed instruction, individualized instruction, and behavior modification. Throughout the chapter, Skinner's ideas are discussed in juxtaposition with other major theoretical ideas in psychology that are currently affecting educational practice.

SKINNER'S SYSTEM OF BEHAVIOR

In 1938, Skinner laid out the fundamentals of his system of behavior in *The Behavior of Organisms: An Experimental Analysis*. This book set forth a formulation of behavior, with accompanying experiments, with which Skinner's later work has been remarkably consistent. Since this book was published, Skinner has extended his analyses into many levels of behavior—animal, human, individual, social—and his work has remained true to the formulations in this early book.

The behavior that Skinner investigated was not unlike that previously studied by Pavlov, Thorndike, and Watson, but there were significant differences. Watson's formulation of behavior, for instance, depended upon a description of a response and the stimulus that elicited it. In contrast, Skinner (1938) found a description of behavior in terms of specific relationships between eliciting stimuli and responses too restrictive and, therefore, inappropriate for describing major classes of behavior. A large body of behavior did not appear to be elicited "in the sense in which a cinder in the eye elicits closure of the lid" (p. 19). Skinner's system consequently makes a significant departure from stimulus-response psychology by concentrating on dual conditioning responses he called *respondent* and *operant*. Respondent behavior is the result of specific eliciting stimuli; operant behavior is emitted by the organism without any specific eliciting stimulus. Changes in frequency of occurrence and rate of performance of the behavior are the dependent variables in Skinner's system.

The principal independent variable of the system is *reinforcement*. In his original work with animals, Skinner showed that as a result of the arrangement of contingencies of reinforcement that are manipulated in precise ways, animals could be trained to perform complex novel sequences of behavior. In Skinner's system, it is the modification of behavior through the appropriate arrangements of environmental contingencies that has been emphasized as a framework for practical implementation. In contrast to experiments that investigate the effects of contrasting conditions in order to prove or disprove a hypothesis, Skinner's experimental tactic has been to introduce environmental conditions that bring about a certain kind of behavior and that maintain the organism's motivation for performing it. The success of Skinner's analytic procedures in leading to highly replicable experiments and the facility with which the techniques involved could

be used appealed to a generation of psychologists and still continue to do so.

Early Counterinfluences

In the late 1940s and early 1950s, in contrast to Skinner, a number of experimenters asserted that an organism can learn through processes of perceptual or cognitive organization that depend only on temporal contiguity of stimulus and response. Those who were proponents of reinforcement as a necessary condition for learning were in the position of defending themselves before these cognitive "latent learning" experimenters. Although the cognitive theorists agreed that reinforcement had an important influence on the modification of behavior, they emphasized that learning and cognitive organization were not satisfactorily explained by the operation of the after-effects of reinforcement (Glaser, 1971).

While there continued to be much debate and experimentation about the effect and interpretation of reinforcement, a new non-Skinnerian interpretation was taking shape. The conception emerging was that reinforcement primarily affects performance and that learning occurs through perceiving relations among events in the environment and storing information about the environment—a process that had been earlier simplistically ascribed to contiguity.

These cognitive conceptions, however, were not as conspicuous as were Skinner's notions at the time. Skinner's exclusive emphasis upon behavioral modification through contingencies of reinforcement and his active disinterest in internal processes and cognitive structures had pragmatic, relevant, and action-oriented appeal.

In 1953, Skinner wrote *Science and Human Behavior*, in which he discussed the implications of his system for humans and the conduct of human affairs. The publication of this text and other activities generated by Skinner and his co-workers marked the strong move of operant psychology into human behavior.

TEACHING MACHINES AND PROGRAMMED INSTRUCTION

Skinner (1954) in an address entitled "The Science of Learning and the Art of Teaching" spoke of translating into educational applications the work done on programming animal behavior. This presentation reflects the force and spirit with which he extended his ideas to the classroom and influenced many at the time who were concerned with the relevance of the science of psychology to education. Skinner was critical of the infrequent use of reinforcement in the

classroom and suggested that a thorough revision of classroom practices was needed based on the advances that had been made in the control of the learning process. He believed that the use of machines for teaching was imperative:

An organism is affected by subtle details of contingencies which are beyond the capacity of the human organism to arrange. Mechanical and electrical devices must be used. . . . The human organism is, if anything, more sensitive to precise contingencies than the other organisms we have studied. (pp. 94-95)

An article that Skinner wrote for a more general audience appeared four years later in 1958 in *Science*. This article, together with the other publications, popularized Skinner's ideas and eventually resulted in the notion of teaching machines and programmed instruction becoming an applied technological enterprise. The article describes the new kinds of machines envisioned by Skinner and the techniques of hinting, prompting, etc., that have become important aspects of programmed instruction.

Interest in the field accelerated rapidly on many levels. Teaching machines of various kinds were sold to schools, industry, and the military. Soon, programmed textbooks on a wide variety of subjects appeared, being more economical to produce than, and conceivably as effective as, simple machines. These texts appeared at all levels of education from elementary schools through professional schools in engineering and medicine.

An expressed concern at this time was that the uncritical acceptance and too-ready use of programmed instruction, coupled with high expectations and awkward application, would result in disappointing outcomes; the rush toward immediate practicality would pull the field away from its loose ties with the scientific study of behavior. As it turned out, things were neither as good nor as bad as expected. The words "teaching machines and programmed learning" were heard less and less, and more was heard about "behavioral science research relevant to education" and educational and teaching technology. Techniques of programmed learning were seen as representing only instances of the ways in which behavioral science might contribute to educational practice (Glaser, 1965).

The ideas of programmed instruction also served to revitalize certain trends in education which had become more or less dormant: the individualization of instruction, the systematic development and evaluation of instructional materials, changes in the nature of testing and grading practices, and the general conception of a discipline of

instructional design. Introducing programmed instruction into a classroom meant changes in the traditional classroom structure: Students would proceed at their own pace and would not keep together as a group; the lecturing skills of the teacher would be deemphasized, at least to the extent that a more tutorial function would be called for; increased time would be given to developing self-study and the self-management of instruction; and, in general, instruction would be closely tied to the progress and accomplishments of each individual.

Also growing out of Skinnerian concepts was a new technique for instruction in higher education. In 1963, Fred Keller, a colleague of Skinner, described a "personalized" course in psychology and, in 1968, published a paper that became the basis for the "Keller Plan." This approach to teaching achieved wide recognition in higher education. Keller's instructional procedure is like programmed instruction in that there is the same stress upon the analysis and mastery of a target performance and on the opportunity for individualized progression and frequent reinforcement, but the sphere of action is different. A student's response is not simply a word or a phrase but rather a verbal explanation of understanding of a principle or use of a technique. Advancing within the program depends less upon the confirmation of small steps than upon a personal interaction between a student and a person with whom there is an interchange of course information. The principles employed in designing this technique for instruction reflect Skinner's assumptions about learning and about empirically determining the effects of environmental conditions in facilitating behavior.

BEHAVIOR MODIFICATION

In the late 1960s work began that led to the application of the principles of operant conditioning to populations in natural settings—non-research-oriented, human-service institutional settings such as mental hospitals, centers for the retarded and those with special behavior problems, and prisons. This work, called behavior modification, received major impetus from Skinner's 1953 book *Science and Human Behavior* and from early research on operant behavior in children and adults with various behavior deviations.

Behavior modification emphasizes environmental, situational, and social determinants as major influences on behavior. The focus is upon behavior that has been learned or needs to be learned, and explicit attempts are made to influence and train behavior rather than alter underlying aspects of personality. Behavior modification relies heavily on Skinner's framework and methods; emphasis is placed

upon the direct observations of behavior and on the events that precede and follow behavior and thereby control it, in the operant sense of control.

The most elaborate behavior modification program undertaken in the 1960s was one reported by Ayllon and Azrin (1965, 1968) in which a token economy was established as a significant aspect of the living environment of patients in a mental hospital. Their demonstration and discussions of the principles involved were seminal, and the idea has been applied in many different institutional settings since that time.

The basic objective of the Ayllon-Azrin undertaking was to design a motivating environment based upon operant reinforcement theory. The building of such an environment would have to consider complex human motivation and the use of reinforcement theory in a complex and more naturalistic environment than had been the case in the past. Ayllon and Azrin describe the general procedures used, the situation and definition of behaviors to be modified, the discovery and use of reinforcing events, procedures for maximizing effectiveness in the award and exchange of tokens, ways of assuring that reinforcement contingencies are in effect, ways of developing behavior when it hardly exists at all, and general administrative and therapeutic considerations in an institutional setting. The procedures were successful in providing a motivating environment for all individuals in the setting regardless of individual characteristics and were effective for individuals of a wide range of ages, education, and years of hospitalization.

As compared with therapeutic applications, applications of behavior modification in educational settings have proceeded relatively slowly; but the work in the schools is accelerating. A review of the research reported in the *Journal of Applied Behavior Analysis* from its beginning in 1968 through the first half of 1976 shows an increasingly large number of studies concerned with the use of behavior modification techniques in classroom settings. However, the bulk of these classroom studies has focused not on enhancing learning or improving academic performance but on decreasing disruptive behavior and its obverse, increasing attention to task. A nondisruptive, hard-at-work class is highly reinforcing to a teacher, and this can account, to some extent, for the emphasis on these variables by behavioral researchers who are faced with the task of convincing teachers that reinforcement strategies are worthwhile. But more important, these variables have been easier to study because the same unit of behavior is observed repeatedly; therefore, measurement procedures directly analogous to those used in the operant laboratory can be applied. In

contrast, learning outcomes that reflect changing knowledge and skill as performance objectives require attention to a variety of additional variables.

A 1974 article entitled "Token Rewards May Lead to Token Learning" (by Levine and Fasnacht) describes studies that question the long-range value of token reward systems. The current cognitive approach of attribution theory suggests that what individuals attribute the cause of their behavior to determines their future behavior. Results obtained from research on attribution indicate that giving a reward can have an effect counter to that desired; giving points, tokens, *M&M's*, etc., to a child for learning a lesson can decrease the intrinsic satisfaction of the lesson. These authors conclude that the operant view of behavior is undoubtedly of value in certain situations and with certain individuals. For example, token reinforcement is especially useful in changing antisocial or inappropriate behavior in children. But tokens were not necessary for the acquisition of these problem behaviors nor for their maintenance, and a first search should be for natural reinforcers of problem behaviors as a significant point of intervention.

A government report on behavior modification issued in 1975 attempted to put the field in perspective through a discussion of what behavior modification is, and what it is not, and through a review of major issues. It states that a range of problems appears to be responsive to behavioral treatment, but it points out that for many types of problems, validating data have yet to be obtained (Stolz, Wienckowski, & Brown, 1975).

COUNTERINFLUENCES

The impact of the application of operant psychology in various forms is widespread. Skinner's work now pervades many settings where the learning and relearning of behavior are significant phenomena—particularly therapeutic situations, institutional environments, and instructional settings at all levels of education. The results obtained are beginning to present evidence with respect to the adequacies and inadequacies of the underlying scientific base. The work of the behavior modifiers has led oftentimes to impressive accomplishments, particularly in circumscribed situations and with the relatively specifiable and less complex aspects of human behavior. However, certain limitations of these applications are beginning to emerge with respect to their long-range generalizable effects and their relevance to much of the internal cognitive world of complex human behavior involved in thinking, problem solving, acquiring understanding of various domains of knowledge, and personal expectations and self-development.

While basic research in operant behavior is being pursued, the present focus of activity is in applied contexts. In contrast, work in the laboratory and in basic theoretical developments is now burgeoning in cognitive psychology, which is oriented toward the complexity of human performance.

Social Learning Theory

The social learning theory of Bandura (1971a) incorporates findings about reinforcement contingencies into a cognitive framework. Social learning theorists accept many of the operational principles of reinforcement and shaping from the work on operant conditioning but interpret the role of reinforcement and learning differently by emphasizing the cognitive symbolic functioning of human beings in acquiring new behavior and in regulating their performance. While Skinnerian theory depicts behavior as the product of directly experienced response consequences, social learning theory emphasizes the fact that humans have the capacity to learn by observation in acquiring behavior without having to build it up by gradual direct trial and error. The learning phenomena described by Skinner that result from direct experience and shaping can also occur vicariously through observation of other people's behavior and its consequences.

In addition, humans are capable of self-regulation. They can manage the environment around them and produce consequences for their own actions in order to control their own behavior. These cognitive and self-regulative influences are important in determining how behavior is acquired and how it is expressed. Such influences upon behavior must be considered in addition to the direct influence of the external environment emphasized by Skinner.

In Bandura's theory, as in the current cognitive psychological theory, a distinction is made between learning and performance. Learning is assumed to occur as a model is observed and symbolic mediators are generated. Reinforcement plays an indirect role largely through the activation of appropriate attentional and mediational processes. Performance, on the other hand, is governed by the laws of reinforcements; the individual is more likely to perform the new behavior when he is reinforced for doing so or when he expects to be so reinforced. Thus, social learning theory does not deny the importance of reinforcement variables in the modification of behavior. It does, however, direct attention to other ways of learning.

Self-reinforcement

The emphasis of social learning theory on human cognitive activity has elaborated the processes involved in self-regulation as exercised by individuals through self-reinforcement (Bandura, 1971b). The fact

that individuals can serve as self-reinforcers and regulate their behavior on the basis of interpretations of self-produced consequences obviously takes on importance in educational settings. The development of the ways in which individuals in school observe their own performance, set criteria for it, and provide themselves with self-reinforcement or self-disparagement is significantly influenced by their classroom experiences with their teachers and peers and by their observations of the modeling influences in these settings.

The investigation of covert self-control is being increasingly pursued, facilitated by the work in cognitive psychology on imagery and other mental processes. The crux of the work is that symbolic outcomes may possess reinforcing properties similar to external physical events. Therefore, an individual, rather than self-administering a tangible reinforcer, can produce a symbolic, implicit consequence that can serve a reinforcing function in regulating overt behavior.

It is conceivable that such concepts of self-regulation and self-control could be taught in school classes as a way of enabling the students to assist themselves in environmental planning and in carrying out behavioral programs that can be used in solving intellectual problems as well as for personal problem solving. At the present time, in programs of individualized instruction combined with open classroom procedures, instruction has been conducted by teaching children a variety of self-management skills in such a way that they can modify the environment for their own learning requirements (Glaser, 1977).

The research done in the context of social learning theory directly confronts the environmental-control aspects of the Skinnerian position. The operant conditioning framework overemphasizes one-way control of the person by the environment and neglects personal control. The work on self-reinforcement and self-control has particularly made us aware that people are able to regulate their own behavior by arranging environmental conditions and administering self-reinforcing consequences.

Human Factors Research and Information Processing

Another counterinfluence, derived from a practical concern with complex human behavior, is human factors research. This research was stimulated by World War II and is concerned with the kind of human performance involved when individuals control complex, man-made systems. Human behavior is studied in terms of its interactions with these systems, and comparisons and analogies are drawn between human processes and the operation of the mechanical systems.

Donald Broadbent, a major influence in cognitive psychology, was one of the central figures in this early research. In a book entitled *Perception and Communication* (1958) he studied attentional processes in terms of the problem of prolonged auditory and visual vigilance involved in tasks such as detecting submarines by radar and flying an aircraft. In this book, models describing the cognitive processes involved in these complex performances are discussed in relation to theories in experimental psychology and an emerging theory of information processing. The effect of these kinds of analyses of human performance was to shift emphasis "from unanalyzed correlation of output and input, towards a consideration of fairly complex internal processes" (Broadbent, 1971, p. 2). Thus, the practical problems of skilled and complex human performance have become a shaping force in current psychological theorizing, in contrast to the simple behaviors studied in operant psychology.

At the present time, however, the major effort in cognitive psychology is largely concerned with developing its theoretical foundations, reminiscent of S-R theory 40 years ago. Cognitive psychology is currently investigating complex human performances such as problem solving, language development, thinking, understanding, attention, memory, imagery, the nature of intelligence, and the problems of comprehending written and auditory discourse. These areas of human endeavor, which have been only peripheral aspects of Skinner's behavioristic psychology, now are the starting points for cognitive psychologists. Realizing the lesson that has been learned—that not only might laboratory work and theory be useful for application, but that application can be a significant test of psychological theory—cognitive psychologists are beginning to take on educational problems.

The combination of the successes and limitations of past theories of behavior with the discoveries and potentials of new theories is what will be shaping the future of educational practice. It can be anticipated that prescriptions for the practical modification of behavior derived from cognitive information processing will agree closely with those of operant conditioning when one is concerned with the training of simple habits in highly restricted situations, such as has been the case in most behavior modification studies in educational settings. "But approaches deriving from the two kinds of theory may be expected to differ increasingly as we turn to problems of education and training of . . . children and adults, where multiple sources of information are the rule rather than the exception and where the individual typically has almost unlimited latitude in selecting the stimuli to which he will respond" (Estes, 1972, p. 729).

It can be further anticipated that educational practitioners will relinquish the older view that human behavior is mainly influenced directly by rewards that strengthen actions, and punishment that weakens them, and begin to design educational settings that are not restricted environments but more natural settings where human cognitive processes take on a more active role. If this takes place, then application of the techniques of operant conditioning derived from the animal laboratory will be less effective in influencing human behavior than applications that take account of the capabilities of the human organism as a processor of information.

CONCLUDING OBSERVATIONS

Skinner identified certain functional regularities of behavior and related them to practical applications. As a good scientist should, he relentlessly pressed for simplifying assumptions and for the full power of his framework. But as happens in scientific progress, constraints have been identified that account for new information, and a deeper and more complete explanation of behavior is being defined.

Skinner's notions of the "environmental space" that dictates human actions appear to come from the adherence of his thinking to the controlled space used in operant work with animals and humans. In situations where the environment can be specified, the behavior of an organism can be predicted and controlled with a high degree of accuracy. The success in the animal laboratory and of behavioral modification in custodial settings provides sufficient evidence. However, when we move out of these boxes and outside of highly restrictive environments, then the nature of control changes to mutual interaction between the activity of organisms and environmental conditions. The individual is now seen as receiving information from the surrounding environment and interpreting this information in order to change and to be changed by the situation. The environment is perceived, processed, and interpreted in terms of states of knowledge in an individual's memory.

The prospects for new directions in education appear to be directly related to revising the concept of a restricted space. Current practical efforts in behavior modification in the classroom appear to demand restricted-space thinking. In contrast, information obtained about cognitive processes should force the design of more malleable learning situations in which the learner acts upon the environment provided, is influenced by it, and in turn changes the situation to further direct his activity.

There is little doubt that the success of operant psychology has been limited to behavior not identified generally with the kinds of intellectual functioning and domains of knowledge to be found in the complex performances of children and adults. Effective theories of learning and instruction require knowledge of what it is that is acquired and of the components and operation of complex intellectual performance. In the classes of performance successfully worked with in behavior modification and programmed instruction, the behavior involved has been relatively easily observed and analyzed. When operant analysis has attempted to identify the nature of complex intellectual performance and the conditions under which it is learned, success has been less evident. It is unclear what the processes and component behaviors are to which the environmental reinforcing contingencies should be applied. A science leading to stronger explanations of complex human performance is necessary to identify what it is that needs to be acquired.

The pragmatic simplicity of the behavioristic, associationistic formulation that attracted the applications of Thorndike and Skinner and that is currently being transmitted to practitioners will undergo change. As complex human activity is being analyzed and better understood, we are becoming increasingly aware of the conditions that foster its acquisition and also impede its development. Current theories of child development and the influence of cultural background, the nature of human problem solving, and the understanding of structures of knowledge all appear at the moment to require constructions that cannot be subsumed under the simplifying assumptions of behavioristic theory.

Cognitive psychology is currently striving for theoretical organization. As the theories that emerge are transferred to frameworks for optimizing human performance that are required for instructional design and theories of teaching, the complexity will be reduced. However, it can reasonably be predicted that instructional design and theories of teaching will be more complex than they are at present and, hence, require more extensively trained educators and more intricate technology for the development of instructional materials and procedures. New ways of fostering learning will emerge based upon a deeper understanding of the skills of reading and language comprehension, the skills of solving mathematical and scientific problems, and the processes that foster intelligence and aptitude. New educational practices will be discovered as a reconstruction takes place involving new knowledge of human cognition and knowledge already accumulated by behavioristic studies of learning.

REFERENCES

- Ayllon, T., & Azrin, N. H. The measurement and reinforcement of behavior of psychotics. *Journal of Experimental Analysis of Behavior*, 1965, **8**, 357-383.
- Ayllon, T., & Azrin, N. *The token economy: A motivational system for therapy and rehabilitation*. New York: Appleton-Century-Crofts, 1968.
- Bandura, A. *Social learning theory*. New York: General Learning Press, 1971. (a)
- Bandura, A. Vicarious and self-reinforcement processes. In R. Glaser (Ed.), *The nature of reinforcement*. New York: Academic Press, 1971. (b)
- Broadbent, D. E. *Perception and communication*. New York: Pergamon Press, 1958.
- Broadbent, D. E. *Decision and stress*. New York: Academic Press, 1971.
- Estes, W. K. Reinforcement in human behavior. *American Scientist*, 1972, **60**, 723-729.
- Glaser, R. (Ed.). *Teaching machines and programmed learning, II: Data and directions*. Washington, D.C.: National Education Association, 1965.
- Glaser, R. (Ed.). *The nature of reinforcement*. New York: Academic Press, 1971.
- Glaser, R. *Adaptive education: Individual diversity and learning*. New York: Holt, Rinehart & Winston, 1977.
- Keller, F. S. "Good-bye, teacher . . ." *Journal of Applied Behavior Analysis*, 1968, **1**, 79-89.
- Levine, F. M., & Fasnacht, G. Token rewards may lead to token learning. *American Psychologist*, 1974, **29**, 816-820.
- Skinner, B. F. *The behavior of organisms: An experimental analysis*. New York: Appleton-Century, 1938.
- Skinner, B. F. *Science and human behavior*. New York: Free Press, 1953.
- Skinner, B. F. The science of learning and the art of teaching. *Harvard Educational Review*, 1954, **24**, 86-97.
- Skinner, B. F. Teaching machines. *Science*, 1958, **128**, 969-977.
- Stolz, S. B., Wienckowski, L. A., & Brown, B. S. Behavior modification: A perspective on critical issues. *American Psychologist*, 1975, **30**, 1027-1048.

Chapter 4 — Summary

THE THEORETICAL IDEAS OF PIAGET AND EDUCATIONAL PRACTICE

GUY J. GROEN

McGill University

ANY account of the relationship between Piagetian theory and educational practice is immediately faced with a number of problems. One is the immense magnitude of the output of Piaget and his colleagues in Geneva. Another is the wide variety of applications reported in the educational literature. Yet another is the fact that Piaget tends to be interpreted very differently by different people, often in a fashion that bears little resemblance to the original.

When I undertook to write this chapter, it seemed that the most reasonable approach would be to describe various applications and trace their relationship to Piagetian theory, drawing a sharp distinction between notions directly derived from Piaget, misunderstandings of Piaget, and influences from non-Piagetian sources. Upon reading through the literature, however, it became clear that this approach would not yield a coherent account of the relationship between Piagetian theory and educational practice, or of the relationships of the various applications to each other. In Piagetian theory, much more than in most psychological theories, everything is related to everything else in an extremely complex fashion. To take something out in isolation is to bias the theory. Moreover, Piaget has written only a few speculative essays on education. They do not form a part of the major thrust of his work. Contrary to the impression conveyed by much of the literature, there is no "méthode de Piaget," no Piagetian theory of education waiting to be applied.

As a result, this chapter is not about applications per se, but about what there is in the theory that is potentially relevant to education.

Specifically, it deals with three issues: first, the nature of the misconceptions that have arisen, especially in North America; second, what the theory actually consists of, and its scientific validity; and third, the educational implications that can be legitimately drawn.

Piaget's career can be divided into three phases, namely, the early work up to about 1930, the period from 1930 to 1957 that saw the development of the major aspects of his theory, and the period since 1957 of his major impact on North America. The mature theory was essentially developed in isolation. For many years, North American psychologists were only familiar with the early work. In the 1950s, this situation changed. An important reason for this was an increased concern with the kind of psychology that might be relevant to education at the high-school and elementary-school levels. It had been kindled by a movement for curriculum reform, begun in the mid-1950s, that received widespread public attention and governmental support after the launching of the Russian satellite Sputnik in 1957.

The leaders of the movement were professional scientists and mathematicians concerned by the fact that the science and mathematics taught in the schools bore little resemblance to what was used by professional engineers, scientists, and mathematicians. Moreover, the conventional curriculum was conveying little of the insight gained by the 19th- and 20th-century mathematics and science. Not only were the schools inadequately preparing future engineers and scientists, but they also were felt to be in danger of producing a generation of scientific illiterates.

The "new curricula" that resulted are so well known that they need not be described here. It suffices to say that the reforms were radical, at least in their original intent. The goal was to create curricula that would teach the average student to understand the basic conceptual structure of the discipline. To achieve this goal, it was felt that the conventional order in which subject matter was taught had to be drastically changed.

The implementation of these reforms raised some very basic psychological issues to which there was no easy way of responding. The new curricula were unique in their emphases on cognitive processes and cognitive structure. Most psychologists were ill-equipped to deal with cognition directly. The educational psychology of the period primarily consisted of applications of other branches of psychology to issues raised by the older, conventional type of curriculum.

It was in this kind of context that the assimilation of Piagetian theory began. There was a search for new approaches, especially a theoretical approach that could account for the development of complex cognitive skills in a way that was responsive to the cognitive issues

raised by the new curricula. Despite the fact that Piaget had said little about education directly, his approach was the only one that seemed to satisfy these specifications at that time.

Conferences began to be held. Probably the first was the 1959 Woods Hole Conference on the improvement of science education.

Books began to appear. There were three that were particularly influential. Bruner (1960) provided a manifesto, loosely based on the proceedings of the Woods Hole Conference, describing how Piagetian theory could resolve the psychological issues raised by the new curricula. Hunt (1961) dealt with the relevance of Piagetian theory to the psychology of mental testing and individual differences, with particular emphasis on the role it could play in a more satisfactory definition of intelligence. Flavell (1963) provided the first thorough guide to the Piagetian literature, and the importance of this guide cannot be overstressed. Whereas Bruner provided a popularization designed to be intelligible to the lay reader and Hunt was only concerned with those aspects relevant to his own theoretical approach, Flavell provided a complete, uncompromising, self-contained account of the entire body of Piaget's work, including much that was not available in translation.

By 1964, the process of rediscovering Piaget was complete. There were now the beginnings of a tradition of Piagetian research in the English-speaking world. However, it differed in important respects from the Geneva tradition. There was a concern with specific educational applications. There was more of a tendency to invent new theories incorporating aspects of Piaget than to develop Piagetian theory *per se*. Most important, there came into existence a theory, usually called Piagetian theory, that was not really Piaget's theory at all but a simplified subset of the theory.

Flavell's book had been too successful. Its treatment of Piaget was so thorough and well organized that it could be used as a primary source, with only occasional references to Piaget's own works to check out the details. For many English-speaking psychologists, Piagetian theory was the theory as described by Flavell rather than by Piaget himself. However, there was a change of emphasis from that of Piaget. Flavell presented the theory rather than explained it. Much of the exposition consisted of direct translations from a bewildering variety of Piaget's writings, with insufficient commentary or clarification. The reader attempting to reconstruct Piagetian theory from Flavell's account was supplied with all the bricks but not enough mortar. The only thing to do was to throw away some of the bricks.

The part of the theory that became accepted in North America, and took on a life of its own, was stage theory. It is probably familiar

to most of my readers. The basic idea is that development is viewed as a sequence of stages: the sensorimotor stage, which ends at about age 2; the preoperational stage, which ends at about age 8; the stage of concrete operations, which ends at the age of puberty; and the stage of formal operations, which characterizes the thought of the adolescent and the mature adult. Each stage is further divided into sub-stages. All this is buttressed by a wealth of demonstrations, involving ingenious tasks designed to show what a child can and cannot do at a given stage or substage.

Stage theory is the villain of our story. It is extremely seductive. It makes sense empirically. It also has very straightforward and obvious educational implications. On the surface, many of the tasks resemble what one encounters at school. It seems eminently reasonable to assume that the stage a child is in has something to do with the way he will perform. Thus, appropriate tests might determine what the child is capable of doing and appropriate training might help him do better sooner. However, there is a trap. Most "applications" of Piaget's theory to education have been naive applications of stage theory more or less in isolation. It will be seen that stage theory, when taken out of its proper context, is a very weak theory indeed. It cannot be applied without major embellishments, which are apt to be highly idiosyncratic and frequently lead to practices that are actually inconsistent with Piaget's own theory! From this stems much of the confusion about Piaget.

To show that this is so, it is necessary to develop a framework for concisely presenting the theory as a whole. In an effort to avoid bias, I adopt a neutral framework, based on the history and philosophy of science. The approach is based on the work of the late Imre Lakatos (1972).

According to this point of view, there is no such thing as a single monolithic theory. What is usually referred to as a theory is really a sequence of theories that change over time. Lakatos calls such a sequence a *research program*. A research program has two components: an invariant *hard core* and a changing *protective covering*. Thus, the component theories differ only in the content of their protective covering. They contain the same hard core, and it is this hard core that characterizes the overall research program.

A research program may be viewed as a vast procedure for generating hypotheses, called *auxiliary hypotheses*, which constitute the constantly changing protective covering. They are generated by two kinds of heuristics: positive and negative. The negative heuristics are designed to protect the invariant hypotheses in the hard core by rendering them incapable of refutation. They respond to anomalies

by changing auxiliary hypotheses in the protective covering.¹ The positive heuristics are there to give coherence to the research program. They form a plan for generating new auxiliary hypotheses and thus elaborating the protective covering, more or less independently of any anomalies that may arise.

A detailed account of the hard core of Piaget's theory, based to a large extent on Piaget's more recent theoretical writings such as *Structuralism* (1968/1971), *Genetic Epistemology* (1971a), or *Biology and Knowledge* (1967/1971), is to be found in the chapter itself. Unfortunately, it is impossible to summarize its contents in a few words that make sense. The important thing is that it consists of propositions about how knowledge is represented by structures, what such structures consist of, and how they develop. The detailed results of stage theory are merely auxiliary hypotheses in the protective covering.

In terms of this analysis, it is possible to be more specific about the dangers involved in taking stage theory out of context. Without the rest of the theory, it is a hollow shell—a set of auxiliary hypotheses without a hard core, or at best a highly impoverished one. The natural tendency is to fill it in, either consciously or unconsciously, with one's own ideas, and the end result is a theory that is no longer Piaget's. Of course, there is nothing wrong in disagreeing with Piaget. The problems arise when the disagreement is not explicitly stated and its consequences are not analyzed in sufficient depth. The result is a new research program with a hard core that is likely to be impoverished and may be based on a set of mutually contradictory hypotheses. The pervasiveness of this practice is one explanation for the existence of such a large number of weak theories, all different yet all attributed to Piaget.

Such misconceived "Piagetian" research programs are very likely to occur in education, and our analysis suggests a logical reason for this. The notion that something will be effective for instructional purposes can be viewed as an auxiliary hypothesis in some kind of research program. However, Piaget's own research program does not contain a theory of instruction. Hence, it is *always* necessary to add to the hypotheses that already exist. Owing to the peculiarly interconnected nature of Piaget's theory, there is a high probability that the investigator who is primarily interested in developing practical techniques with a short-term payoff will develop a set of auxiliary hypotheses that are inconsistent with the hard core.

¹ Such a change may take a number of forms. It may be a major modification of the theory, it may simply be an ad hoc way of rendering the anomaly harmless, or it may constitute an attack on a rival research program.

To discover whether anything at all is consistent with the hard core, it is best to look closely at what Piaget himself has had to say about education.

Piaget has not written any books on education per se. What he has done, over the years, is to write a series of essays. These are of two kinds. Some deal in a statesmanlike manner with very general policy issues only indirectly related to the theory. They need not concern us here. The others deal more specifically with the relationship of the theory to practical educational matters. However, the treatment is always at a quite general level, and specific applications are almost invariably only mentioned in passing. Two of these essays are particularly important.

The earliest is *The New Methods: Their Psychological Foundations*, which was written in 1935. It appears in Piaget (1971b). The new methods are those that were used in the progressive education of the 1930s, which emphasized "learning by doing" or fostering activity in the child. Piaget is in favor of progressive education. In the traditional school, the child's activity is inseparable from a continual constraint exercised by the teacher. The progressive school, on the contrary, appeals to real activity, to spontaneous work based on need and interest. However, there is a problem. In the traditional system, a child's behavior is to a great extent predictable and controllable by the system itself. If, on the other hand, use is made of "progressive" methods where a child is free to adapt to his environment, then a much more precise knowledge of the child is needed in order to understand how and when they are effective.

Taken as a whole, the essay is an exposition of principles that might be applicable, not a description of how to apply them. However, there are two very important hints. The first is the notion that play is crucial in ensuring the assimilation of reality. The second is a warning against too great a neglect of the environment, especially its more structured aspects. Thus are introduced two themes that play a crucial, and complementary, role in the success or failure of any attempt to apply Piagetian theory. Free activity is important, and so is the structure of the environment. Yet the more explicit a structure we put on the environment, the less opportunity a child has for free activity, and vice versa. The critical issue in any Piagetian application is: How does one achieve a satisfactory balance between these two apparent opposites?

The second essay is *Education and Teaching Since 1935*. It, too, appears in Piaget (1971b) but was written in 1965. Intended as a sequel to the first essay, it is addressed to the question of why there exists no "science of education sufficiently developed to be able to

answer the innumerable questions that present themselves every day and whose solution is therefore left either to empiricism or to tradition."

Much of Piaget's analysis in this essay deals with administrative and political problems. However, he does devote a considerable amount of attention to one issue that is closely related to direct applications of the theory. This is the tendency for educational research to exist in a state of isolation from scientific disciplines that could support it with the principles it needs. Thus, educational research may have quite powerful and sophisticated evaluation techniques available, but the ability to explain the results is almost totally lacking. To Piaget, only an interdisciplinary approach can provide such explanations. In particular, it is obvious, to Piaget, that educational research will have to employ a precise psychology, not merely that of common sense.

Piaget proceeds to outline how his own approach can contribute to a precise psychology of this kind. He begins by restating the basic problem of the 1935 essay in a more mature form. There are "traditional" methods of education, based on the notion of gradually incorporating children into adult social life and determined by adult norms. There are "progressive" methods, which emphasize the spontaneous and relatively autonomous aspects of the child's intellectual structures. Presumably, these are determined more by what a child can actually do well. However, the distinction is no longer a pejorative one. To Piaget, the traditional methods are perfectly good for teaching things such as spelling that do not have any real epistemological base but are simply things that are arbitrarily defined by the norms of adult society. The issue is whether there might not be better methods for teaching mathematics and science and other areas that have developed through a process of research and discovery. To Piaget, such subjects can be taught in two ways: passively, as an assemblage of facts and rules, or actively, by means of some process that involves the reconstruction or reinvention of facts and rules. He immediately makes it clear that he prefers the second way on ideological grounds. The question he is interested in dealing with, however, is whether there exist good underlying psychological reasons for this preference.

Piaget proceeds to indicate how his research program provides some such reasons. The argument is complex. Its gist is as follows. The principal aim of education is to develop intelligence. The essential functions of intelligence consist of understanding and inventing. In terms of the hard core of the research program, understanding a phenomenon consists of generating transformations belonging to a structure that models it. Invention consists of developing the structure. In much the same way (it would seem) as assimilation and

accommodation are complementary, so are understanding and invention, and intelligence itself is a condition of equilibrium that results from their functioning. Empirically, the development of intelligence corresponds to the sequence of stages provided by stage theory. Theoretically, it corresponds to the development of structures with an ever-increasing capacity to assimilate and organize (or model) reality. This development depends on natural, or spontaneous, processes whose existence constitutes a necessary condition for the effectiveness of any form of instruction. They may be utilized and accelerated by education, but they are not derived from it. However, Piaget once again takes pains to warn us that the fact that these processes are natural and spontaneous does not imply that they are determined solely by maturation. In fact, little is known about the role of maturation, according to Piaget, except that it exists. It can do no more than open up possibilities. The theory is about how such possibilities become actualized.

This brings us to the important point. It follows from Piagetian theory that these natural and spontaneous processes are analogous to the methods of research used in developing science and mathematics. Indeed, the child can be viewed as carrying out his own research program and using it to construct his own sequence of theories about reality. That is how he or she assimilates knowledge. There are, of course, other things involved in the acquisition of knowledge. In particular, knowledge can be communicated through books, speech, pictures, symbols, or demonstrations. Hence, traditional methods have their place. However, structures must be developed to assimilate or model this knowledge and this can only be done through mental or physical activity involving these natural and spontaneous processes. Hence, the curriculum should make room for such activities. Most important, the opportunity for these activities must be extensively provided in science and mathematics because it is only through the performance of such activities that these subjects can be properly understood.

Unfortunately, most practical "applications" of Piaget's theory pay little attention to applying these notions. Instead, they are direct applications of stage theory with very little else added. That such projects may fail has no bearing on the adequacy of Piaget's own notions. Indeed, Piaget's complete theory would predict that many such applications are a waste of time.

It is the structures that are of primary importance, rather than the stages and the tasks used to define them. A test result indicating that a child is "in" a stage only has significance in that it indicates the minimal structures available. The fact that some training procedure

improves performance says little, of and by itself, unless an attempt is made to discover what has happened from a structural point of view. Moreover, training studies may have little practical educational significance if they focus exclusively on the auxiliary hypotheses of stage theory (in much the same way as a project designed to increase performance on a subtest of the Stanford-Binet might lack educational significance). The stages were designed to characterize intelligence rather than knowledge. The acquisition of knowledge involves other cognitive functions, and their development may follow different patterns of stages. It is the interaction of all these functions, and not intelligence per se, that is likely to determine how well a child learns in a practical educational setting.

Some clues as to how Piaget's notions might be legitimately used in educational practice are to be found in the work of Kamii (Kamii & DeVries, 1977) and Papert (e.g., Papert, 1972). Kamii's primary focus is on teacher training, at the preschool level. It has three important aspects. The first is the notion that Piaget's theory offers a tool rather than a method or set of principles that can be directly applied. However, it is not a tool that exists ready-made nor can it be constructed by the developer of a prescriptive curriculum. It is a tool that must be constructed by classroom teachers themselves. The second is the notion that any curriculum that prescribes fixed tasks is doomed to be inconsistent with Piagetian theory. What must be done instead is to create an environment in which the teacher or the child can generate tasks as the need and opportunity arise. The third is that the main goal, as far as the child is concerned, is to enhance the formation of the structures that will ultimately characterize formal operational thinking.

Viewed objectively, Kamii's approach is not necessarily any better than the more direct or prescriptive uses of Piagetian theory that we have considered up to now. What is important, for our purposes, is that it demonstrates that it is possible to use Piaget's ideas, in practical educational settings in a nontrivial way, and remain completely consistent with the theory. It is unclear whether Kamii's approach is the only way to accomplish this, but it does suggest a paradigm for achieving consistency. Also, it suggests that it is a paradigm that cannot be followed without paying a price: It involves abandoning the notion of the structured curriculum as a direct instrument of intervention in the teaching process.

Papert's computer-based LOGO environment gives an unusually concrete embodiment to Piaget's rather abstract notions of what is involved in inventing and developing new structures. The instructional approach is to follow Piaget's model of the child as a researcher

and the teacher as research director. The child selects a project and is free to do anything he wishes to accomplish it, subject to quite explicit constraints imposed by the LOGO environment. The important thing is that it is a project in which the child has to worry about the procedures he or she is using. A plan must be conceived, executed, and debugged.² Debugging is particularly important. The environment ensures that a child will obtain automatic and quite subtle feedback regarding the errors in his or her current plan. Successful debugging can be accomplished only by deliberately reflecting upon what the plan really consists of and what there is about it that is causing the errors.

What is important about all this is that it follows from Piagetian theory that the child who has invented the computer program must have, in his or her head, some structure that models the *functioning* of the computer program. As he or she goes through the debugging process, this structure must change in the same way as the program changes. Thus, the child's structures will accommodate and assimilate reality as a result of the debugging process. Moreover, the child is much more likely than usual to be aware of the assimilation and the accommodation that are actually going on. In the long run, it seems reasonable to assert that the child will develop general heuristics for organizing programs so that it is easy to accomplish desirable goals, and so that the programs themselves are easier to debug. These general heuristics (Papert calls them "powerful ideas") are at the heart of LOGO. The basic hypothesis is that they are not simply ways of organizing computer programs. They can be utilized to organize all kinds of procedural knowledge. They are also at the heart of Piagetian theory. They must reflect the functioning of very high level logico-mathematical structures. According to Piaget, it is the existence of such structures that is responsible, in the long run, for the higher levels of intelligent human activity.

A psychological theory should serve education as a source of good ideas. Does Piaget's theory provide any good ideas? Enough has been said to indicate that there may be some. However, enough has also been said to indicate that finding them is likely to be a long and arduous process. Unfortunately, the deceptively simple ideas that result from a superficial interpretation of Piaget are almost all bad. What is needed is much greater sophistication on the part of researchers interested in the educational implications of Piaget's theory. What may also be necessary is a greater recognition of the fact that

² A technical term in computer programming for the iterative process of discovering why a program does not work. Anybody who has ever written a nontrivial computer program should be familiar with the peculiarly self-motivating nature of this process.

there are serious gaps in Piaget's theory, especially where issues of instruction are concerned. It may well be that a considerable amount of basic research is necessary to fill in those gaps. In the meantime, it might lessen the current state of confusion if developers of educational products were more willing to acknowledge that they were filling the gaps with their own ideas, and refrain from the currently fashionable practice of attributing them to Piaget.

REFERENCES

- Bruner, J. S. *Process of education*. New York: Vintage Books, 1960.
- Flavell, J. H. *The developmental psychology of Jean Piaget*. Princeton, N.J.: Van Nostrand, 1963.
- Hunt, J. M. *Intelligence and experience*. New York: Ronald Press, 1961.
- Kamii, C., & DeVries, R. Piaget for early education. In M. C. Day & R. K. Parker (Eds.), *Preschool in action* (2nd ed.). Boston: Allyn & Bacon, 1977.
- Lakatos, I. Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge*. Cambridge, England: Cambridge University Press, 1972.
- Papert, S. Teaching children thinking. *Programmed Learning and Educational Technology*, 1972, 9(5), 245-255.
- Piaget, J. *Genetic epistemology*. New York: Norton, 1971. (a)
- Piaget, J. *Science of education and the psychology of the child*. New York: Viking, 1971. (b)
- Piaget, J. *Biology and knowledge*. Chicago: University of Chicago Press, 1971. (Originally published, 1967.)
- Piaget, J. *Structuralism* (C. Maschler, trans.). New York: Basic Books, 1971. (Originally published, 1968.)

Chapter 5 — Summary

PSYCHOANALYSIS AND AMERICAN ELEMENTARY EDUCATION

PATRICK SUPPES AND HERMINE WARREN

*Institute for Mathematical Studies
in the Social Sciences
Stanford University*

IN ASSESSING the impact that psychoanalytic theory has had on American elementary education, we limit ourselves to the education of normal children as well as to influence in the United States. We begin at the turn of the century, when references to psychoanalysis began to appear in the American literature, and end in 1957. The year 1957 marks the symbolic death of the Progressive Education movement as well as the launching of Sputnik and on both these counts is broadly thought of as an end of an era in education.

PSYCHOANALYTIC THEORY PERTINENT TO EDUCATION

In this section we sample some of the literature relevant to education that has received considerable attention in the psychoanalytic literature. We begin with a précis of Anna Freud's *Psychoanalysis for Parents and Teachers* (1935), in which her famous lectures to teachers are to be found.

Anna Freud's Lectures to Teachers

The first two lectures concern the prelatency child, i.e., the child before he enters school at the age of 5 or 6, and are devoted to the development of the following themes. First, one must consider education to begin on the first day of life, and not when the child enters the

formal school setting. Second, one must take the instinctual life of the child into account if education is to succeed, or, put another way, too narrow a view of the purpose of education is certain to fail. Education, whether administered by the parents or by the teacher, has as its general aim to make the child become like the adults around him, and the achievement of this purpose requires a constant struggle with the child's instincts. Miss Freud lays out some of the main components of this instinctual life, emphasizing early sexual development, sibling rivalry, the Oedipus complex, and the transference of this constellation of family feelings (at least in muted form) to the school.

The third lecture concerns the latency period, which is marked by the repression of the instincts that are characteristic of the earlier stages. That the child is especially receptive to learning during the latency period has been recognized from time immemorial, and hence traditional schooling is begun for the child during just this receptive period. Adolescence, long regarded as marking the beginning of sexual feeling, in fact only witnesses a repetition of the earlier stages prior to the latency period, and the school must be prepared for this change.

Almost uniquely among important philosophies of education, psychoanalysis recognizes an explicit need to limit the goals of education. Without such limitations, education will tend to overly restrict the important part that should be played in life by the instinctual drives. In the psychoanalytic view, the profound problem of education is to find the right proportion of instinct gratification and instinct restriction.

Theory of Learning

Because the psychoanalytic view toward learning emphasizes increasing differentiation of structure and capacity for sublimation, it is not surprising to find that a characteristic theme of considerable significance in psychoanalytic approaches to learning is the emphasis on readiness for learning or preconditions for learning. Psychoanalysts are undoubtedly right in saying that classical learning theories in academic psychology have not properly emphasized the deeper reasons for readiness for learning. The postulated stages of development, leading into the latency period around the age of 5 or 6 years, provide a conceptually compelling explanation of why this is the appropriate time to begin serious school learning—more so than is provided by any competing theory known to us despite the difficulties of formulating the theory in suitably precise form.

Other articles are much concerned with reward and punishment, or what in learning-theoretic terms would often be called positive and

negative reinforcement. In an influential article on the means of education, Fenichel (1945) summarizes matters this way: "And now it seems clear that we have identified the three basic means of all education, namely, direct threat, mobilization of the fear of losing love, and the promise of special rewards. The second means can be applied effectively only if the child has previously experienced the fear of losing love" (p. 285). Ideas of a similar nature, though not as systematically worked out, are presented in Peller (1946). She expresses what has come to the fore in recent theories of learning, namely, the importance of postulating internal structures. The theories of learning she criticizes certainly were not powerful enough to generate or account for the structures characteristic of the complex thinking and behavior of the child. It is characteristic of a psychoanalytic viewpoint toward these matters to insist that neither internal structure nor environmental conditions alone can provide an adequate account of complex development.

The last group of articles we sample focuses on the topic of curriculum and instruction. Zachry (1933) discusses the role of instruction in art for personality development and presents a case for the inclusion of art in the curriculum. Bettelheim (1948) discusses the ways in which the social studies teacher can help students master the problems of adolescence. He emphasizes that real-world mastery of his problems is not possible for the adolescent, so that classroom analysis of these problems is especially important.

Learning Disabilities

In the early psychoanalytic view, learning inhibitions result from unstable sublimations of instinctual drives. The most conspicuous manifestation of intellectual inhibition during the early years of school life is in reading difficulty, and an excellent systematic survey of psychoanalytic contributions to the problem of reading disabilities is to be found in Blanchard (1946). Blanchard presents a survey of the statistical data on sources of reading difficulty, and, referring to Gates' (1923/1930) extensive study, accepts his estimate that emotional disturbances are present in about 75% of children with serious reading disabilities, that in about a fourth of these cases the emotional difficulties are the cause of the reading difficulty, and that in the remaining three-fourths of the cases the causal direction is reversed.

A survey of the literature on emotional components of learning arithmetic is given in Plank and Plank (1954). Their method of study consists in the examination of autobiographies for remarks about the pleasure or pain of learning mathematics. Though their methodology

is flawed, their results support the claims of several writers, e.g., Levy (1943) and E. Klein (1949), that overprotected children who are strongly attached to their mothers often find learning to read easy and learning arithmetic difficult.

There are other instances in which psychoanalysts have brought a clinical-case approach to bear on the issue of how affective factors may interfere with cognitive processes. Melanie Klein (1924) gives a rich sample of such clinical analyses. Fantasies about particular letters of the alphabet and particular numbers are perhaps the most familiar examples.

The articles cited above discuss instances in which the child's ego is overwhelmed by other forces. Pearson (1952) considers learning disabilities as a problem in ego psychology as such. Pearson discusses at some length the child for whom learning is boring because it requires effort and interferes with immediate pleasures. Pearson is raising an issue that has increasingly engaged the attention of educators and psychoanalysts, that of the child whose ego development has been retarded by too permissive an upbringing.

Discipline and Moral Development

To a large extent, the psychoanalytic view of learning and moral development is related to the theory of learning outlined above, since psychoanalytic writings on these topics are more concerned with the promotion of inner control than with overt behavior.

In general, the psychoanalytic position on discipline and moral development has been misunderstood. A common popular image is that psychoanalysis has encouraged, as has no other movement, the undisciplined child who is permitted to develop in his own way and without the external constraint that he conform to the attitudes and habits of adults around him. This widely held view is a caricature of what psychoanalysts actually write about the problem of discipline. Indeed, psychoanalytic writings on this subject stress the need for avoiding any simple approach to these problems. On the one hand, impulses must be allowed to surface and to be expressed; on the other, restraints on the child's instinctual demands must be imposed or his ego development will suffer. These ideas are spelled out with particular clarity in articles by Sperling (1951) and Kris (1948, 1950).

The Teacher

One of the more insightful discussions of the teacher from a psychoanalytic standpoint is to be found in Bernfeld (1925/1973), who was one of the prominent psychoanalytic pedagogues in Vienna

just after World War I. Bernfeld points out the psychological dangers inherent in the teaching situation and comments on the ordinary teacher's too frequent lack of insight into this problem.

Other authors address the problem of how to differentiate the role of the teacher from that of the therapist. A British publication by Searl (1932) draws a number of useful contrasts between psychoanalysis and education. Searl emphasizes the usefulness of a personal psychoanalysis for teachers but at the same time expresses the limitations of what might be expected. An article appearing at about the same time, much referred to later, is the one by Erikson (then Homburger; 1935). In the opening passages of this article he contrasts the roles of the teacher and the therapist by pointing out that the analyst is required to remain for the most part a silent observer while the teacher's work involves continual talking. Without offering any panacea for these complicated problems, Erikson concludes on the optimistic note that self-knowledge can only lead to more effective teaching and that a personal psychoanalysis for the teacher should provide a powerful means for its achievement.

EVIDENCE OF IMPACT

There has not been a psychoanalytic movement in American education, but other movements have, to a greater or lesser extent, carried psychoanalytic ideas into the school. We consider here, in succession, child study and child development, educational psychology, mental hygiene, progressive education, and the general culture.

The Child-study and Child-development Movements

One measure of the impact of a theory on a profession is the extent to which it is represented in the standard textbooks. For this reason, we looked at all the more frequently cited texts in child study and child development used in the training of teachers for the years 1903-1956 to determine the number of pages giving either explicit, implicit, or derived coverage to psychoanalysis. *Explicit* coverage refers to pages on which psychoanalysis, Freud, or some other orthodox psychoanalyst is mentioned; *implicit* coverage refers to pages not included in explicit coverage but on which psychoanalytic concepts like that of the defense mechanisms are mentioned; and *derived* coverage refers to pages that mention neo-Freudians (like Karen Horney or Harry Stack Sullivan), projective tests, and experimental attempts to test psychoanalytic ideas as well as attempts to combine psychoanalytic approaches with other approaches (e.g., Dollard & Miller, 1950). We found that all three kinds of influence increase over time, even though the average number of pages remains relatively small. Implicit

impact was greater than explicit impact, and derived impact was greatest of all. The greater increase in the coverage of derived topics probably reflects the closer relation between projective tests and similar topics, on one hand, and the primarily behavioral orientation of child development during this period, on the other.

The limited representation of psychoanalysis in the texts can be partly explained by the fact that all the texts cover an enormous amount of material and a wide range of topics, and throughout most of the period covered by our review the authors are highly restrained about adopting a particular theoretical viewpoint of their own. But it is also the case that many of the authors did not feel at home with the clinical traditions out of which psychoanalysis came, and it is our general impression that they did not wish to become very involved in an exposition of it and often proceeded with an obviously ambivalent attitude.

Educational Psychology

As in the case of the child development movement, we examined the outstanding textbooks used by educational psychologists for evidence of explicit, implicit, and derived impact of psychoanalysis (as already defined above). Again, the data show that influence increases with time, but to a lesser extent, and that the average number of pages remains quite small. Undoubtedly, the theory of the defense mechanisms received the most coverage under the heading of implicit impact, although the authors referred as frequently to mechanisms not discussed within the psychoanalytic literature as they did to the mechanisms discussed first by Freud and later by his daughter Anna in her classic *The Ego and the Mechanisms of Defense* (A. Freud, 1936/1946), as well as others. For example, alongside brief discussions of rationalization, projection, reaction formation, and so forth, we find mentioned such mechanisms as "logic-tight compartments," the "polyanna" syndrome, and the "sour grapes" syndrome, suggesting that consideration of the defense mechanisms was facilitated by their similarity to ideas coming from other sources.

Representative sample remarks about Freud made by the textbook authors suggest that dislike of the sexual emphasis of psychoanalysis, as well as prevailing learning-theoretic conceptions of behavior, foreclosed the possibility of explicitly accepting psychoanalytic theory. Throughout, the texts were deficient in giving the proper attribution to Freud as the source of the mechanisms, a situation Hilgard (1949) has described well in his presidential address to the American Psychological Association. However, in surveying 83 English-language educational psychology textbooks by decades, Gates, Jersild, McKillop, Rivlin, Shoben, and Watson (1956) report a steady and

significant increase in the percentage of texts covering the topics of mental hygiene, unconscious motivation, and counseling, guidance, and psychotherapy, as well as a uniform increase in the number of pages devoted to these topics. As their own extensive references make apparent, it is reasonable to attribute some of this development to psychoanalysis.

It is probable that much of the changed emphases in the texts can be attributed to student and teacher pressure. As early as 1911, O'Shea reported in the second volume of the *Journal of Educational Psychology* that a survey of graduate students in education showed that a majority thought "physics was as close to the problems of training children as this work in educational psychology" (p. 331). A similar and highly influential article was published later by Watson (1926), who sampled some faculty members at Teachers College, a group of experienced teachers and administrators, and a group of undergraduates who had not taught but were preparing for teaching, concerning what should be taught in educational psychology. Watson concludes that "the apparent concurrence of opinion would place major emphasis upon problems of emotion and personality adjustment, problems of original nature and heredity, and problems of general teaching method. Little emphasis would be placed upon the psychology of the acquisition of skill, or the psychological problems involved in home-making and pre-school child training, or the various schools and theories in psychology" (p. 596). Two years later in a survey of the opinion at a teacher-training institution concerning the relative value of subject matter in educational psychology to the elementary-school teacher, Hertzberg (1928) found that, in agreement with Watson's survey, problems concerning personality adjustment were felt to be a subject of major importance.

Mental Hygiene

Indicative of the impact of psychoanalysis on the mental hygiene movement are the various studies of teacher attitudes made during this period. Using more than 500 teachers as respondents, Wickman (1928) conducted a classic investigation that revealed marked differences between the the interpretation of children's behavior problems by teachers and by mental hygienists. Although teachers rated as most serious transgressions against authority, dishonesty, and sexual immoralities, the mental hygienists rated these items as relatively minor in importance but as more serious unsocial, withdrawing, or regressive behavior. The hygienists also rated as fairly serious cruelty, temper tantrums, and truancy, each of which was not so rated by the teachers.

Hunter (1957) undertook a detailed study to compare classroom teachers in 1955 with those studied by Wickman in 1926-1927. One of the more striking changes he found was in the ranking of sexual activity. Teachers in the Wickman study ranked heterosexual activity and masturbation first and third in seriousness; the teachers in Hunter's study moved these ranks to 19th and 28th, respectively. More generally, Wickman reported a correlation coefficient of $-.11$ between the rank order of traits in terms of seriousness by the mental hygienists and by teachers. Using the same mental hygienists' ranking that Wickman did, he was able to report a correlation coefficient of $.22$ between the teachers' responses and the hygienists'. We attribute some of the marked changes in attitude toward sexual behavior to the indirect influence of psychoanalytic ideas as they increasingly permeated the culture during the period covered by this review.

These findings are broadly consistent with several studies that attempted to assess how much teachers know about mental hygiene. An excellent example is a study by Hobson (1937), who conducted an extensive survey of teachers in North Dakota organized along the following lines. First, a list of mental hygiene principles was constructed. The author began with a list of 160 principles and sent it to 12 individuals working in mental hygiene, so as to derive from the 160 principles a reliable subset of 42. A simple imaginary case exemplifying each of the 42 principles was constructed, and judgments about some aspect of the case were given as test items to teachers. Among the 42 principles, we would identify five or six as having possible antecedents in psychoanalytic theory, although the language of psychoanalysis was not used explicitly and it would be reasonable to claim other theoretical sources as responsible for the inclusion of these principles. An example of such a principle is "The adult's reaction to the child's natural curiosity about sex may result in having the child carry over after a sense of guilt that attaches to anything that has to do with sex" (pp. 233-236).

We examined 22 texts used in courses on mental hygiene. The trends are the same as those for child development and educational psychology, but the much greater average number of pages showing the explicit impact of psychoanalysis on mental hygiene is clear evidence of greater impact. As would be expected, the general content of the texts on mental hygiene is considerably closer to psychoanalysis than the other texts discussed above, but, as is also evident, some topics judged important on the basis of their coverage are quite separate from psychoanalytic concerns.

Despite the increasing acceptance of Freudian theory, it remains the case that the discussions of the defense mechanisms are usually

separated almost completely from those of psychoanalysis as such. Since many of these same texts refer frequently to a group of texts and treatises written in the earlier part of the century (roughly in the period 1911-1918), we investigated the analysis of the defense mechanisms in these earlier books and the extent to which the authors attributed the source of the concepts to Freud. We restrict our comments here to the books by one of the most prominent people in this period, William A. White, who was probably the most important medical figure in the early history of mental hygiene.

White published three widely read books (1911, 1916, 1917). All three explicitly acknowledge Freud as the source of the concepts of defense mechanisms. The 1916 book, for example, has as its subtitle *An Introduction to Psychoanalysis*.

Progressive Education

In general, the progressive education movement was not as concerned with the psychology of the child as with practical problems of school reform. Cremin (1961, pp. 306-308) gives a summary of the main impact of the progressive education movement on the mainstream of American education, and of his ten summary points, none primarily concerns the psychology of the child. On the other hand, the move away from the traditional formal curriculum, the introduction of variation and flexibility in the grouping of students, and the introduction of new curriculum materials are features of change that will be permanently associated with the movement.

Still, the progressive movement is popularly associated with the introduction of the concept of the child-centered school, and in many cases Sigmund Freud and other psychoanalysts were cited in support of this view. All in all, however, there was not much discussion of psychoanalysis in the progressive literature, and the views that lay back of the child-centered school, though they overlapped with many of the concepts of psychoanalysis, were by no means identical. The psychological points of disagreement are these.

1. That struggle and conflict are required in order for secondary processes to assume dominance over the more primitive primary processes was not seen as essential to learning, although from a psychoanalytic viewpoint the existence of this conflict is fundamental.

2. The limitations of what the teacher can accomplish and the inevitable dangers of the power placed in the teacher's hands, as emphasized by Bernfeld and Erikson, for example, received little if any recognition in the progressive education movement. All that was needed, it was thought, was the reform of the teachers' attitudes so as to move from old to new methods of instruction. The fundamental

psychological problems that exist in such relationships as surrogates for earlier family relationships were scarcely recognized.

3. In the case of discipline, there was certainly recognition by many progressive educators that the naive concept of absolute freedom for the child was not workable, but the deeper psychoanalytic insight that as the child matures the free instinctual forces must be repressed was scarcely acknowledged by many prominent progressive educators.

4. Concerning learning difficulties, the unconscious roots of such difficulties in some children were on occasion explicitly acknowledged by progressive educators, but their criticisms were so overwhelmingly directed at the formal methods of the past that the origin of learning difficulties in the internal life of the child was seldom stressed.

Cultural Impact

A large number of historians of art, literature, and society have analyzed and commented on the remarkable influence Freud has had on American culture in the 20th century (e.g., Sullivan, 1932; Hofstadter, 1955; Hoffman, 1957; Hale, 1971). Our perusal of magazines and journals read by the educated public produced substantial evidence of awareness of Freudian theory.

Another source of change in thinking came through the parent-education movement. The most important American child-rearing manual of the 20th century, namely, Benjamin Spock's *Baby and Child Care*, is close in conceptual formulation to psychoanalytic thought. Both Sulman (1973) and Bach (1974) have argued by the use of contextual content analysis that Spock generally accepts the psychoanalytic view of child development in terms of infantile sexuality, Oedipal conflict, and the active role of the unconscious, even though the more technical psychoanalytic concepts are not presented either in standard psychoanalytic terms or in commonsense substitutes.

CONCLUSION

In the preceding pages we have attempted to present the part of psychoanalytic theory that is most relevant to the education of normal children and to assess its impact on aspects of American elementary education. In this final section we summarize our conclusions.

The Diffuse Influence of Psychoanalysis on Education

Psychoanalysis has had a significant impact on 20th-century elementary education. The most salient psychoanalytic ideas that

were incorporated into educational thought were the importance of the early years of childhood, the fact of childhood sexuality, and the role the defense mechanisms play in internal life.

Although it seems reasonable to attribute to psychoanalysis much of the changed approach to young children in school, its explicit influence in the professional educational literature during the first half of this century was not substantial. Psychoanalysis developed in a medical setting and its history is evident in the main lines of its influence on education, so that only the mental hygiene movement showed a substantial explicit influence of psychoanalytic ideas. This literature in itself reflects a profound change in attitudes toward education, however, since no longer are the central aims of education solely the acquisition of basic skills and a fund of knowledge but now include the development of a healthy and effective personality.

The Increasing Influence of Psychoanalysis over Time

In broad terms, the influence of psychoanalysis in education grew during the entire first half of this century. The fact that psychoanalytic ideas were incorporated into education more and more over this period is in its way evidence of the magnitude and depth of their impact. It is also evidence of how slow the application of new scientific ideas can be. It took a long and complicated route for the impact of Freud to be felt and for the ideas to become part of our thinking about the child and therefore about the education of children.

Ambivalence Toward Psychoanalysis

Ambivalence toward psychoanalysis was exhibited in a number of ways. Textbook authors we have examined from various fields of education more often than not adopted such an attitude in their explicit comments on psychoanalysis, and in the extensive presentations of the defense mechanisms it was uncommon for their origin in psychoanalytic thought to be given proper attribution. Although the almost uniform theoretical eclecticism of the textbook authors reflected an ambivalence toward theory in general, this was especially the case toward a theory whose ideas and methods ran contrary to the prevailing academic views and procedures.

Nonprofessional Channels

To a very considerable extent the channels through which the influence of psychoanalysis spread were not professional ones. It seems fair to say that the dissemination of psychoanalytic ideas through the public press and through the very large literature addressed to parents had as much influence as the professional literature in education itself. Teachers' attitudes changed on a variety of

topics in a direction more consistent with psychoanalytic concepts over most of the period we have surveyed, and it seems likely that much of the basis for this change was the dissemination of psychoanalytic ideas throughout the general culture rather than through any particular professional literature.

Limitations of Education

In contrast to an emphasis on unlimited growth and the unbounded potentialities of education characteristic of the writings of John Dewey and, more generally, of the progressive education movement, psychoanalytic thinking about education, whether in the home or in the school, has stressed the natural limitations of what can be achieved and what should be expected. The view of limitations ranges from accepting the necessary existence of conflict in the developing child to recognizing that teachers are subject to the same kinds of psychological problems in relating to children as parents are in relating to their children. This strong sense of the limitations of what education can hope to achieve was not a psychoanalytic theme that was widely accepted during the period we review, although it is a salient feature of the intellectual climate of today.

REFERENCES

- Bach, W. G. The influence of psychoanalytic thought on Benjamin Spock's *Baby and Child Care*. *Journal of the History of the Behavioral Sciences*, 1974, 10(1), 91-94.
- Bernfeld, S. B. *Sisyphus or The limits of education* (F. Lilge, trans.). Berkeley: University of California Press, 1973. (Originally published, 1925.)
- Bettelheim, B. The social-studies teacher and the emotional needs of adolescents. *School Review*, 1948, 56, 585-592.
- Blanchard, P. Psychoanalytic contributions to the problems of reading disabilities. *Psychoanalytic Study of the Child*, 1946, 2, 163-187.
- Cremin, L. A. *The transformation of the school: Progressivism in American education, 1876-1957*. New York: Knopf, 1961.
- Dollard, J., & Miller, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
- Fenichel, O. The means of education. *Psychoanalytic Study of the Child*, 1945, 1, 281-292.
- Freud, A. *Psychoanalysis for teachers and parents* (B. Low, trans.). Boston: Emerson Books, 1935.
- Freud, A. *The ego and the mechanisms of defense*. New York: International Universities Press, 1946. (Originally published, 1936.)
- Gates, A. I. *Psychology for students of education* (Rev. ed.). New York: Macmillan, 1930. (Originally published, 1923.)
- Gates, A. I., Jersild, A. T., McKillop, A. S., Rivlin, H. N., Shoben, E. J., Jr., & Watson, G. Educational psychology. *Review of Educational Research*, 1956, 26(3), 241-267.

- Hale, N. G., Jr. *Freud and the Americans: The beginnings of psychoanalysis in the United States, 1876-1917*. New York: Oxford University Press, 1971.
- Hertzberg, O. E. The opinion of a teacher-training institution concerning the relative value of subject-matter in educational psychology to the elementary school teacher. *Journal of Educational Psychology*, 1928, **19**, 329-342.
- Hilgard, E. R. Human motives and the concept of self. *American Psychologist*, 1949, **4**, 374-382.
- Hobson, C. V. How much do teachers know about mental hygiene? *Mental Hygiene*, 1937, **21**, 231-242.
- Hoffman, F. J. *Freudianism and the literary mind* (2nd ed.). New York: Grove Press, 1957.
- Hofstadter, R. *The age of reform: From Bryan to F.D.R.* New York: Knopf, 1955.
- Homburger, E. Psychoanalysis and the future of education. *Psychoanalytic Quarterly*, 1935, **4**, 50-68.
- Hunter, E. C. Changes in teachers' attitudes toward children's behavior over the last thirty years. *Mental Hygiene*, 1957, **41**, 3-11.
- Klein, E. Psychoanalytic aspects of school problems. *Psychoanalytic Study of the Child*, 1949, **3-4**, 369-391.
- Klein, M. The role of the school in the libidinal development of the child. *International Journal of Psycho-Analysis*, 1924, **5**, 312-331.
- Kris, E. On psychoanalysis and education. *American Journal of Orthopsychiatry*, 1948, **18**, 622-635.
- Kris, E. Notes on the development and on some current problems of psychoanalytic child psychology. *Psychoanalytic Study of the Child*, 1950, **5**, 24-46.
- Levy, D. M. *Maternal overprotection*. New York: Columbia University Press, 1943.
- O'Shea, M. V. Psychology in the normal school. *Journal of Educational Psychology*, 1911, **2**, 327-333.
- Pearson, G. H. J. A survey of learning difficulties in children. *Psychoanalytic Study of the Child*, 1952, **7**, 322-386.
- Peller, L. E. Incentives to development and means of early education. *Psychoanalytic Study of the Child*, 1946, **2**, 397-415.
- Plank, E. N., & Plank, R. Emotional components in arithmetical learning as seen through autobiographies. *Psychoanalytic Study of the Child*, 1954, **9**, 274-293.
- Searl, M. N. Some contrasted aspects of psycho-analysis and education. *British Journal of Educational Psychology*, 1932, **2**, 276-296.
- Sperling, M. Psychoanalytic aspects of discipline. *Nervous Child*, 1951, **9**, 174-186.
- Spock, B. *The commonsense book of baby and child care*. New York: Duell, Sloan & Pearce, 1946.
- Sullivan, M. *Our times: The United States 1900-1925* (Vol. 4). New York: Scribner's, 1932.
- Sulman, A. M. The humanization of the American child: Benjamin Spock as a popularizer of psychoanalytic thought. *Journal of the History of the Behavioral Sciences*, 1973, **9**, 258-265.

- Watson, G. What shall be taught in educational psychology? *Journal of Educational Psychology*, 1926, **17**, 577-599.
- White, W. A. *Mental mechanisms*. New York: Journal of Nervous and Mental Disease Publishing Company, 1911.
- White, W. A. *Mechanisms of character formation*. New York: Macmillan, 1916.
- White, W. A. *The principles of mental hygiene*. New York: Macmillan, 1917.
- Wickman, E. K. *Children's behavior and teachers' attitudes*. New York: Commonwealth Fund, 1928.
- Zachry, C. B. The role of art in personality development. *Mental Hygiene*, 1933, **17**, 51-58.

Chapter 6 — Summary

THE IMPACT OF THE LANGUAGE SCIENCES ON SECOND-LANGUAGE EDUCATION

H. H. STERN, MARJORIE BINGHAM WESCHE,
AND BIRGIT HARLEY

*Modern Language Centre
Ontario Institute for Studies in Education*

THE impact of scientific research on educational activity can be attributed in part to the influence of great men of science. Such men leave their mark on a discipline and, as a result of the importance of that discipline for education, may transform educational theory or practice in a decisive way. Men like Dewey, Thorndike, Freud, and Piaget have had this type of deep-going influence on educational thought in the 20th century.

At other times it is a dominant scientific concept, technique, or theory that has produced an effect upon education, as with the concept of intelligence, the notion of the intelligence quotient, and the mental-testing movement.

Then again, there is a third way in which scientific developments can influence educational practice, illustrated in the present case study of the impact of research on language teaching. There is no one scientist nor one scientific concept that can be singled out as having exercised a *decisive* influence on language teaching. Yet over the past hundred years and particularly in recent decades, the developing language sciences have had an immense effect on all phases of language teaching. It is no longer possible to think about language pedagogy or to make decisions in this field without taking note of their pervasive influence. This is not to say that there are no great figures. Among earlier linguists, Sweet, Palmer, and Jespersen stand out as pioneers; in more recent decades, Bloomfield, Fries, and Chomsky among linguists, and Carroll among psychologists, can be

singled out in this way. Nor can it be denied that specific language-related theories, techniques, and concepts have made their way into language teaching: the phonetic alphabet of the 1880s, the concept of linguistic structure in the 1940s, the notions of "habit" and "skill" in the 1950s, "competence," "performance," and "creativity" in the 1960s, and more recently "communication." Throughout the hundred-year period reviewed in this case study, however, it is much more the influence of the language sciences as a whole than the impact of a single figure or concept that has been felt in language teaching.

THE IMPACT OF LINGUISTIC THEORY

Theoretical linguistics has benefited language teaching by creating among language teachers an awareness of language beyond the particular language taught. There is today a much clearer understanding than there used to be of what it is we are doing when we teach or test a language. There were of course exceptions in the earlier stages of language-teaching theory. Gouin and Palmer, for example, had highly developed theories of language well before linguistics had evolved as a modern science, but their theories were of an *ad hoc* nature. In the modern era, language teachers have a body of theory to refer to. On the whole, language pedagogy has attempted to absorb the developing linguistic theories into its own theories for teaching language. Consequently, in more recent writings on language teaching and in course materials, we have observed the shift from traditional grammar to structural linguistics, and more recently to transformational grammar and, in present-day thought, to the rising interest in semantics and sociolinguistics.

The constant changes in theoretical interests, emphases, and formulations have, however, raised questions in the minds of both language teachers and theoretical linguists. How far can language teaching be guided by linguistic theory? To what extent do these changes make more than a superficial contribution to language-teaching practice? Because of the violent controversies in linguistic theory during the 1960s, this problem has been discussed repeatedly between 1966 and the early 1970s. In general, the advice that language teachers have been given is to adopt an independent stance towards linguistic theory. This is much more easily said than done, because in order to choose from the many ideas in linguistics and psycholinguistics, the teacher must have a much more wide-ranging and sophisticated knowledge of the language sciences than it is fair to expect.

Another problem can be illustrated by sociolinguistics, which, as a

growing field of study, is transforming our view of language and in principle is highly relevant to language teaching. However, theoretical insights outdistance our capabilities to turn them into practical teaching ideas. As so often happens, linguists have developed new concepts and formulated a new model (such as the model of the speech act), but there are, as yet, little data to support it, let alone data for the commonly taught languages. Nonetheless, we are already clamoring for teaching materials embodying these new ideas and modifications in procedures.

Here lurks the danger that language teaching has run into again and again in its relationship to theoretical linguistics (as well as to psychology). There is a shift in theoretical outlook, perfectly legitimate in the development of scientific theory, and to accommodate it—or rather, to keep up with the Joneses—language pedagogy makes slight adjustments in its techniques. A few new terms are sprinkled across the old fare without any fundamental rethinking of the relative importance of the new concepts to the established practice.

On the positive side, however, linguistic theory challenges our view of language and can, in the long run, have a profound and beneficial influence on curriculum content and the treatment of language in language teaching.

THE IMPACT OF DESCRIPTIVE LINGUISTICS

The importance of descriptive linguistics for language education was fully realized by linguists of the 1940s. For example, much current teaching of English as a second language is still based on the research done by Fries and Lado in the 1940s and 1950s. Contrastive analysis was an attempt to bring descriptive linguistics even closer to the needs of language teaching. To be sure, descriptive and contrastive linguistics have suffered something of an eclipse as a result of the preoccupation with theory building between 1965 and 1970, but language teaching cannot do without descriptive investigations if it is to remain up to date, authentic, and reliable. Fortunately, some linguists have not been overly influenced by the changing theoretical emphases and have been prepared to continue a task in which they believed. Otherwise we would not have such recent works of descriptive scholarship as *A Grammar of Contemporary English* (Quirk, Greenbaum, Leech, & Svartvik, 1970). Description is particularly necessary today, even in the languages most widely taught, because contemporary views of language are broadening and include an understanding of sociolinguistic and cultural aspects. The fact that this kind of

research is in short supply illustrates the problem that priorities in linguistics do not necessarily coincide with those of language teaching.

Several applied linguists around the world have tried to come to grips with the further problem of making language descriptions accessible and useful for language teaching. Some have proposed a data bank of linguistic information so that the course writer or teacher would have a source of reliable and up-to-date linguistic information constantly at his disposal. The data, of course, cannot remain in a data bank alone; hence the idea of a "pedagogical grammar" and of a "sociolinguistic and sociocultural guide," which would serve as a pool of ordered information on language, society, and culture and be the necessary intermediary between theoretical and descriptive research and curriculum development.

THE IMPACT OF PSYCHOLOGY AND PSYCHOLINGUISTICS

The influence of psychology and psycholinguistics has taken a different direction from that of linguistics. While in the linguistic field new concepts have been formulated without specific regard for second-language teaching and have then been adopted or tried out in language teaching, psychological research has been more responsive to specific problems in education and has sought to confirm and refine notions commonly accepted in the language classroom. Thus the ideas of a "talent for languages" or a "will to learn" were not invented by psychologists but have been defined more precisely in terms of the concepts of aptitude and motivation and can now be related in a more conscious way to other factors in the psychology of language learning and of learning in general. The earlier failure of psychology, and to a certain extent psycholinguistics, to recognize the process of second-language learning as a problem of theoretical interest has, until recently, held up progress in the development of a theory of second-language learning. In this context, it bears note that psychologists did so little for so long to modify the oversimplified view of language learning that has been widespread in language pedagogy. This deficiency is now being remedied.

Among the important contributions of psychology to language teaching are the following: (a) The understanding of second-language skills (comprehension and production) has been deepened by psycholinguistic analyses; (b) the study of first-language acquisition has stimulated empirical investigation of whether it can serve as a model for second-language learning; (c) an important issue now receiving attention is whether there are developmental sequences in

first- and second-language learning to which language teaching should be geared; (d) certain characteristics of the learner—aptitude, motivation, and individual learning style—have been investigated, and their relationships to each other and to ultimate success in language learning are now better understood; and (e) recent studies of the learning process itself can be expected to lead to important new insights.

On the other hand, some of the most vital issues remain unresolved: (a) The issue of optimal age is still controversial; (b) the widespread failure to learn a second language, although empirically obvious, remains theoretically as puzzling as the inevitability of first-language acquisition; (c) questions about individual differences in language learning and language-learner characteristics are still open; and (d) the psychological bases of many second-language teaching techniques are not clear.

Overall, the contact with psychology and psycholinguistics has contributed concepts, models, and basic theoretical positions to the teaching of language; it has defined important questions and has suggested possible interpretations. The total volume of psychological and psycholinguistic studies has indirectly created a greater sensitivity to various psychological factors in language learning. At the same time, it clearly shows that the psychological problems of second-language learning demand direct study. They cannot be resolved by extrapolation from other psychological studies.

THE IMPACT OF EDUCATIONAL RESEARCH

A last major way in which research has had an impact on language teaching is through studies in institutional settings. These studies have demonstrated that research is not necessarily done elsewhere and then "applied" but that it can deal directly with the issues of language teaching in the classroom. In all, these studies have yielded a much more realistic understanding of language teaching.¹ A far less dogmatic view of teaching methodology has spread. One would be hard put today to claim that any narrow set of techniques, specific technological invention, or particular starting age is the "open sesame" of language teaching. In spite of the violent controversies that have surrounded such research, the sobering findings of studies

¹ Outstanding examples of the value of research in school settings are the IEA studies of English and French as second languages (Lewis & Massad, 1975; Carroll, 1975).

such as the Pennsylvania Project or the Keating Report have largely been confirmed by the test of time. It is hardly conceivable that major changes in language education could occur in the future without taking account of research.

ASSESSMENT OF MECHANISMS

The language sciences focus on language and language behavior in general; that is their strength. But from the point of view of language pedagogy, it is also a weakness, for it is in the nature of language teaching to be concerned with a particular language. Thus the problem is that those who are concerned with language in general do not themselves teach languages, while those who teach individual languages lack the general theoretical background that would make them more responsive to language-teaching research needs.

It is our contention that, in order for language research to produce a substantial effect on language teaching, it must eventually capture the interest of those directly concerned with the teaching of particular languages in schools and universities. Retrospectively, it appears to us that the single most important cause of the discrepancy between language-research needs and the systematic provision for research and training is the failure of modern language departments in universities throughout the world to have radically revised their role in language research and language education.

Where, then, did the initiative come from for the relevant research done so far? Governments, charitable foundations, university departments of psychology and education, and newly created departments of linguistics have, to a large extent, displayed the initiative that language departments, with a few exceptions, appear to have lacked. But governments are subject to political pressure; their language interest may rise and fall in unpredictable ways, and research funding, while it has often been unexpectedly generous, may become too erratic to allow for the continuing growth that is needed. There may, for example, be a preference for funding projects that have high "visibility" or that promise to produce "instant" results, rather than studies that fit into a planned framework that takes several years to develop. For a psychology or linguistics department, second-language learning can only be a partial interest if it is not to unbalance seriously their general terms of reference. Under these circumstances, it is surprising how much has so far been achieved in different places around the world in building up a research effort and theoretical discussions of considerable scope.

CONCLUSION

Has the association of language teaching and research led to more effective language learning? A very critical view of the relations between research and practice indicates very little concrete evidence that would associate specific improvements in language learning with specific results of research. If a comparison is made between the effects of language research and the effects of agricultural, pharmaceutical, or industrial research, it is found that the evidence equivalent to an increased crop yield, a complete medical cure, or heightened industrial productivity is lacking in language teaching. The easiest excuse to give would be that the input-output model used in other kinds of applied research is not applicable. It is, indeed, perhaps unjustified, as Macnamara (1974) has pointed out, to judge the results of language research with the yardstick of commercial productivity.² Nevertheless, we want to be assured that there have been true benefits from the research activity, even if they cannot be measured accurately. This case study would seem to indicate that the impact of language-related research has generally been positive, but that the dramatic advances in language teaching which some had hoped for have not materialized.

After events have taken place, one is inclined to take them for granted. Thus, we do not question enough the prolonged interaction between language research and language pedagogy. In some ways, it is surprising that the impact has been as great as it has. What, then, brought language teaching and research together? Research has not come into language teaching as the result of an overwhelming clamor from language teachers. The language teaching profession has, on the whole, long been aware of the difficulties and failures in its endeavors, but in seeking solutions its tendency, until recent decades, has been to speculate, to rely on personal opinion and experience, to try out new methods and institute reforms mainly on a trial-and-error basis, and to resolve ideological conflicts by polemic and debate. Nor did the development of linguistics point inevitably to language teaching as an obvious field of application. Between the 1930s and 1950s, linguistics, as a relatively new science, was more concerned with establishing itself as an autonomous field of study than it was with applications. However, when in the early 1940s the need arose for practical

² As Macnamara (1974) puts it: "Somehow education in the U.S. has fallen into the hands of businessmen who demand that inputs, processes and outputs be specified in quantified terms" (p. 47).

applications of linguistics to language teaching, linguists were eager to demonstrate the relevance of their discipline. Returning to the question of what has brought language teaching and research together, we believe it is largely the failure of language teaching to solve its persistent problems in any other way. And it is the equally persistent need, and demand, to come to grips with these problems that has brought about an even closer and more regular association between language research and second-language teaching.

In order to benefit from a science, the educator must regularly, and critically, scan and review the science in question. The scientist who claims to influence the educator with his scientific theories and research findings must try to understand the educational activity in its full complexity, as it occurs. However, the relationship is less productive if the science imposes itself on the educational activity, or, similarly, if its dictates are uncritically accepted by the educator. Briefly, the conclusion reached is that the relation between research and practice is most fruitful when (a) the scientist recognizes the full scope of the educational activity in its own right, (b) the educator, in turn, recognizes the scope of the scientific discipline in its own right, and (c) the relation is seen as a dialogue between educator and scientist and not as a monologue by either.

REFERENCES

- Carroll, J. B. *The teaching of French as a foreign language in eight countries*. New York: Wiley, 1975.
- Lewis, E. G., & Massad, C. E. *English as a foreign language in ten countries*. New York: Wiley, 1975.
- Macnamara, J. What can we expect from a bilingual program? *Working Papers on Bilingualism*, 1974, 4, 42-56.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. *A grammar of contemporary English*. London: Longmans, 1970.

Chapter 7 — Summary

PARADIGM AND PRACTICE: ON THE IMPACT OF BASIC RESEARCH IN EDUCATION

J. W. GETZELS

University of Chicago

THE skepticism of educators about the bearing of research—and particularly of basic or theory-oriented research—on the operation of schools is quite extraordinary. Even more extraordinary is the skepticism of the researchers themselves about the bearing of what they are doing on the operation of the schools. On the grounds of assertions that research findings have little practical bearing on the actual operation of schools, support for research in education by foundation and government agencies has been sacrificed for other, more “applied” activities, which presumably have instant practical bearing. I shall argue that the substance of the assertions that research has little bearing on the operation of schools and the patterns of support for research that follows from them are in egregious error; both must be altered if advances are to be made not only in educational inquiry but also in educational practice.

My argument proceeds in four parts. First, basic or theory-oriented research can have powerful effects on the operation of schools; I would challenge those who favor support for applied activity at the sacrifice of such studies to match the effects of their work against the effects that have accrued from the studies I shall cite here. Second, far from contributing “only tiny bits” of fragmentary and irrelevant knowledge, basic or theory-oriented research contributes to the broad conceptions and fundamental paradigms of human behavior; thus it provides the crucial contexts and guides for the substance and method of education. Third, these conceptions and paradigms enter into the preparation of school personnel and, in fact,

as the paradigms are altered, preparation is altered (despite assertions to the contrary, ideas do have consequences). Fourth, I show the crucial practical, and ultimately applied, function the university performs through its basic research in formulating and pursuing apparently impractical problems: the definition of problems that might otherwise remain unpursued for being unformulated. It is this role that is being enfeebled by the unhappy confluence of financial stringency and the anti-intellectualism of funding agencies, which reflects their mistaken beliefs regarding the separation of basic research from practical application. And finally, I comment briefly on two competing views of the relation of theory, research, and practice and indicate the consequences of each view in education.

EFFECTS ON SPECIFIC CLASSROOM PRACTICES

Despite the belief that basic research has little effect on practice, the fact is that its effects can be powerful, as evidenced by the specific basic studies mentioned below which have had manifest effects on our attitudes about pupils, leadership style, classroom management, and the nature of the curriculum.

In the case of giftedness, for example, it was long held that genius and insanity are related, and, unless a child's intellectual precocity were restrained, the child would come to a bad end since "early ripe, early rot." There was generally a belief that intellectual giftedness was inseparable from personal deficiency. Terman's basic research (Terman, 1925; Burks, Jensen, & Terman, 1930; Terman & Oden, 1947, 1959) showed that, instead of being inferior, the gifted children were superior to other children not only intellectually but also physically, socially, emotionally, and morally. Moreover, their superior achievement in school applied so generally to the different school subjects as to refute the traditional belief that gifted children are usually "one-sided." Terman's periodic follow-up of the original gifted children showed that they maintained and even increased their early superiority. As Carroll (1963) pointed out, the attitudes we hold today toward giftedness stem from Terman's basic research.

Another instance of basic research that had far-reaching effect on practice was an apparently simple little study by Lewin, Lippitt, and White (1939) in which they observed that children's behavior appeared to be a function of whether the leader was "autocratic," "democratic," or "laissez-faire." Within less than a decade after the study was reported, scores of textbooks and school programs applied the findings to education. There was hardly an educator who was not made aware of the study and of the effects of these "styles of leadership" on the "group dynamics" of the classroom and the school.

According to the doctrine of formal discipline, just as muscular powers could be strengthened through arduous physical exercises like running, so the mind could be strengthened through educational exercises—exercises that need have no knowledge value in themselves. This doctrine was used to justify Latin and Greek in the curriculum, since these proved manifestly arduous exercises. Thorndike and Woodworth (1901a, 1901b, 1901c) tested the influences of exercises in technique, devoid of content, on one's ability to handle other problems. The general conclusion was that, while there was some improvement, it was not due to any "mysterious transfer of practice" but rather to a functioning of identical elements in the practice series and in the final test series. The Thorndike and Woodworth experiments were widely cited by the curriculum makers and contributed to profound alterations in the course of studies in the school.

EFFECTS ON BROAD EDUCATIONAL POLICY

The second issue is the charge that, even when it is productive, basic research provides only "tiny bits of knowledge" and, as such, is without influence on the broad problems of education. It is precisely basic study in its cumulative impact that alters the general conceptions, and ultimately what Kuhn (1962) calls the paradigms, of the human being and his condition, which in turn have an influence on the broadest educational problems.

Consider, as one example, the alterations in architecture of the classroom, symbolic of more profound but less visible changes in our conceptions of the "learner."

Within many urban areas one can find four types of school buildings of varying age and classroom architecture. In one type of building, dating from the turn of the century, the spaces called classrooms are rectangular, the pupils' chairs are bolted to the floor, and the teacher's desk is front and center. In a second type of building, dating from the 1930s, the classrooms are square, the pupils' chairs are movable, and the teacher's desk is out of the way in a corner. In a third type of building, dating from the 1950s, the classrooms are also square but the pupils' desks are trapezoidal in shape so that when placed next to each other they form a circle, and the teacher's desk has vanished. In a fourth type of building, there is a classroom constructed in the 1970s that is four times the size of the ordinary classroom and has no teacher's or pupils' desks but is filled instead with all manner of odds and ends from finger paints to Cuisinaire rods. For purposes of reference, the spaces in the first type of building will be called "rectangular" classrooms; in the second, "square"

classrooms; in the third, "circular" classrooms; and in the fourth, "open" classrooms.

At the turn of the century the psychological conception of the learner was that he was more or less an ideationally empty organism associating stimuli and responses through the operation of rewards and punishments under the control of the teacher. In effect, learning was "teacher oriented." It was no accident, then, that the prevailing methods and materials and even the school buildings and classrooms were teacher centered. Typically, the classroom was rectangular with the teacher up front and the pupils in chairs rigidly fastened to the floor in straight rows facing forward so that they could not turn away from the presumed source of their learning experience: the teacher.

It became evident that this earlier conception did not account for all or even most of the behavior observed in the learning situation (Segal & Lockman, 1972). The individual did not seem to experience what he was to learn as discrete stimuli. Instead he saw apparently discrete stimuli as "belonging together," as making a configuration or gestalt. In the learning laboratory, experimenters were dealing with problems of "perception" and "prägnanz"; and teachers in the classroom, with problems of "insight" and "closure." Moreover, personality theory was postulating that the human being was a bundle of needs, interests, values, persuasions, and other psychic forces that determined what the learner would and would not respond to.

The conception of the learner as an empty organism was transformed into that of the learner as an active organism. From this point of view the learner—not the teacher—was the center of the learning process. It was no accident, then, that the teacher-oriented classroom became the pupil-centered classroom. The teacher's desk was shifted out of the way and the pupils' rigid desks were replaced by movable chairs that could be shifted at will according to the need of the pupils.

Drawing from the emerging social-psychological and field-theory views, the child as learner began to be envisioned as not just an active organism but a social organism, and learning was perceived as occurring in a field of interpersonal actions, each person serving as a stimulus for every other person. It is hard to overestimate the impact on the classroom of the leadership and group climate studies by Lewin and his associates beginning in the late 1930s. As a result of basic studies such as these, learning became more "group centered." If learning itself is a social process, then a circular or group-centered classroom where everyone faces everyone else is the most sensible and practical learning environment.

There has, however, been a growing discontent with the paradigm of behavior that underlies the views of the learner represented by the

three classroom architectures just mentioned. They were founded in a combination of the homeostatic model of self-maintenance and the drive- or tension-reduction theory of behavior (Getzels, 1964). Learning, thinking, problem solving, and intellectual exploration may be means of reducing certain drives. But they may also be ends in themselves, the organism acting to *increase* as well as to decrease stimulation.

Piaget (1952, 1954), Hunt (1960), Schachtel (1959), and White (1959), among others, point out that the central fact in the development of children is not hunger satisfaction or thirst satisfaction or some other primary drive satisfaction; it is the opportunity for interaction with the environment as manifested in the child's curiosity and exploratory activity. From this point of view, the learner is not only a problem-solving and stimulus-reducing organism but also a problem-finding and stimulus-seeking organism. The altered classroom arrangement that seems to be emerging from the conception of the learner as a stimulus-seeking organism may be characterized as "inquiry centered."

Conceptions of the learner—conceptions having had their source, at least in part, in basic research—do have an impact on what goes on in the classroom. Far from providing only bits of knowledge, basic research affects our paradigms of human behavior and thus influences the forms and practices of the classroom.

EFFECTS ON PROFESSIONAL PREPARATION IN EDUCATION

Basic research and paradigms have an impact not only on practice but also on professional preparation for the schools. In the case of educational administrators, a comparison of the textbooks indicates the influence of research. Between the foreword to the first edition of perhaps the most widely used text in educational administration (Campbell, Corbally, & Ramseyer, 1958) and the foreword to the fourth edition a dozen years later there is a crucial change—from reliance only on the *experiences* of already practicing administrators to the study also of the *concepts* of administration.

During the first quarter of this century, emphasis was on "scientific management." Management consisted of "knowing exactly what you wanted men to do, and then seeing to it they do it in the best and cheapest way" (Taylor, 1911, p. 21). Numerous treatises and textbooks—those by Bobbitt, Cubberley, Strayer, and Reeder—applied the principles to educational administration.

What came to be called the "human relations" view developed

during the second quarter of the century. This view held that human relations were the essence of an organization and that everyone in it should "evoke each other's ideas . . . and integrate their viewpoints in pursuit of the common goal" (Metcalf & Urwick, 1940, p. 14). Support for the human relations view came from studies by Elton Mayo and his colleagues, notably F. J. Roethlisberger and William J. Dickson (1939). Educational administration was further influenced by Lewin's inquiry into different styles of leadership. The primary responsibility of the educational administrator came to be seen as facilitating interaction in the faculty group so that "all individuals affected by any decision should have a share in determining its character and form" (Yauch, 1949, p. 16).

Although the principles of scientific management and those of human relations were often antithetical, they had in common the conception of administration as a vocation or trade, not as a field of study. About mid-century a different view arose that synthesized the two—the administrator not only must know what to do but also must understand why he is doing it. The central issue in administration became the synthesis of the personal or individual and the impersonal or organizational requirements in order to maximize the "actualization" of both.

This conceptual shift moved the focus in the preparation of administrators from a *solution orientation*, in which experienced administrators gave prescriptive answers to how-to-do-it questions, to an *inquiry orientation*, in which problems were posed. There was a concomitant shift in the research during this period. Virtually none of the texts before 1950 referred to systematic theory of administration; virtually none of these after 1960 did not refer to it. Most telling in this respect is to follow the introduction of new ideas from the research reports through successive editions of the same authors' texts for preparing school administrators.

BASIC RESEARCH AND THE UNIVERSITY

Asked of any other institution, What is it about? What does it do? the answer is relatively simple. The work of the university and its scholarship lies uniquely in making explicit that which is problematical, in exploring enigmas that lack formulation as problems, to say nothing of solutions. As Einstein put the issue, "The formulation of a problem is often more essential than its solution. . . . To raise new questions . . . requires imagination and marks real advance in science" (Einstein & Infeld, 1947, p. 95) and, it may be safe to say, in activity of any worth. This, then, may be the paradox: Research may have the

greatest effects on education not so much where it attempts to alter an element of practice directly, as indirectly—where it raises new questions and contributes to transformations in the general paradigms of the human being.

THEORY, RESEARCH, AND PRACTICE

The contention of this chapter is not that basic research inevitably has practical uses, or that the practical uses of the basic research have invariably been salutary. But in light of the increasing number of assertions that basic research is of little value for education, coupled with the increasing sacrifice of support for basic research for presumably more practical activities, it seems necessary to consider in specific terms the view that even for the curriculum, structure, and administration of schools there may ultimately be nothing so practical as research that is "merely theoretical." As Dewey (1929) said,

Theory is in the end . . . the most practical of all things, because the widening of the range of attention beyond nearby purpose and desire eventually results in the creation of farther and wider-reaching purposes. (p. 17)

REFERENCES

- Burks, B. S., Jensen, D. W., & Terman, L. M. *Genetic studies of genius: The promise of youth* (Vol. 3). Stanford, Calif.: Stanford University Press, 1930.
- Campbell, R. F., Corbally, J. E., & Ramseyer, J. A. *Introduction to educational administration*. Boston: Allyn & Bacon, 1958.
- Carroll, J. B. *Educational psychology: Achievements and prospects*. Paper presented at the meeting of the National Society of College Teachers of Education, Chicago, Feb. 14, 1963.
- Dewey, J. *Sources of a science of education*. New York: Liveright, 1929.
- Einstein, A., & Infeld, L. *The evolution of physics*. Cambridge, England: Cambridge University Press, 1947.
- Getzels, J. W. Creative thinking, problem-solving, and instruction. In E. R. Hilgard (Ed.), *Theories of learning and instruction* (Sixty-third Yearbook of the National Society for the Study of Education, Part 1). Chicago: University of Chicago Press, 1964.
- Hunt, J. M. Experience and development of motivation: Some reinterpretations. *Child Development*, 1960, **31**, 489-504.
- Kuhn, T. S. *The structure of scientific revolutions*. Chicago: University of Chicago Press, 1962.
- Lewin, K., Lippitt, R., & White, R. K. Patterns of aggressive behavior in experimentally created "social climates." *Journal of Social Psychology*, 1939, **10**, 271-299.
- Metcalf, H. C., & Urwick, L. *The collected papers of Mary Parker Follett*. New York: Harper & Row, 1940.

- Piaget, J. *The origins of intelligence in children*. New York: International Universities Press, 1952.
- Piaget, J. *The construction of reality in the child*. New York: Ballantine, 1954.
- Roethlisberger, F. J., & Dickson, W. J. *Management and the worker*. Cambridge, Mass.: Harvard University Press, 1939.
- Schachtel, E. G. *Metamorphosis*. New York: Basic Books, 1959.
- Segal, E. M., & Lockman, R. Complex behavior or higher mental processes: Is there a paradigm shift? *American Psychologist*, 1972, **27**, 46-55.
- Taylor, F. W. *Shop management*. New York: Harper & Row, 1911.
- Terman, L. M. *Genetic studies of genius. Mental and physical traits of a thousand gifted children* (Vol. 1). Stanford, Calif.: Stanford University Press, 1925.
- Terman, L. M., & Oden, M. H. *Genetic studies of genius. The gifted child grows up: Twenty-five years' follow-up of a superior group* (Vol. 4). Stanford, Calif.: Stanford University Press, 1947.
- Terman, L. M., & Oden, M. H. *Genetic studies of genius. The gifted group at mid-life: Thirty-five years' follow-up of the superior child* (Vol. 5). Stanford, Calif.: Stanford University Press, 1959.
- Thorndike, E. L., & Woodworth, R. S. The influence of improvement in one mental function upon the efficiency of other functions (Part 1). *Psychological Review*, 1901, **8**, 247-261. (a)
- Thorndike, E. L., & Woodworth, R. S. The influence of improvement in one mental function upon the efficiency of other functions (Part 2). *Psychological Review*, 1901, **8**, 384-395. (b)
- Thorndike, E. L., & Woodworth, R. S. The influence of improvement in one mental function upon the efficiency of other functions (Part 3). *Psychological Review*, 1901, **8**, 553-564. (c)
- White, R. W. Motivation reconsidered: The concept of competence. *Psychological Review*, 1959, **66**, 297-333.
- Yauch, W. A. *Improving human relations in school administration*. New York: Harper & Row, 1949.

Chapter 8 — Summary

EDUCATIONAL RESEARCH AND EDUCATIONAL REFORM: A CASE STUDY OF SWEDEN

TORSTEN HUSÉN

*Institute for the Study
of International Problems in Education
University of Stockholm*

THE FRAMEWORK FOR POLICYMAKING

THE extensive inquiries as well as planning activities preceding national legislation in education and affecting the entire system of education in Sweden are to a large extent carried out by ad hoc committees, so-called Royal Commissions, which constitute a unique element in Swedish policymaking. When fundamental public policy issues emerge, such as basic principles of income taxation, provisions of medical services, or structure of the school system, pressures begin to build for the setting up of a governmental committee to take a deep look into the problem area. The responsible cabinet member brings the matter up with his colleagues and eventually gets the approval of the cabinet to appoint a Royal Commission to conduct a comprehensive inquiry into the matter. The cabinet also agrees on rather detailed terms of reference for the commission, which in some cases even prescribe the policy according to which the recommendations are to be made. But in most cases the commissions are free to conduct their inquiries completely unbiased.

In selecting the members of a commission, the government tries to see to it that the various interest groups are represented in a reasonably balanced way, for example, members of Parliament from various political parties and representatives of central federations of trade unions and other potential pressure groups. A commission operates a secretariat with access to the expertise of specialists who can be called

upon to serve on the commission. Since the 1940s, these governmental committees have increasingly commissioned research focused on central problems within the domain under scrutiny. Thus, several key commissions, whose recommendations since 1945 have been instrumental in the reform of the Swedish educational system, have initiated some of the most important research projects in education conducted in the country.

The very *modus operandi* of a commission has often been conducive to a rather close liaison between research workers and policymakers as seen in the following example, which the author himself had an opportunity to experience.

The 1957 Commission provided the basis for the 1962 Education Act, which made provision for the introduction of comprehensive education nationwide in Sweden after a 10-year period of pilot programs. The commission sponsored an extensive 3-year survey of the school system of the city of Stockholm to assess the outcomes of the introduction of the comprehensive school structure. The design of the survey as well as its progress was discussed regularly over a 3-year period in the full commission, which was presided over by the Minister of Education. Therefore, without delving too deep into technicalities, policymakers on the commission had ample opportunity to familiarize themselves with research strategies and methodologies employed by the educational research workers as well as with their theoretical and interpretative frameworks.

The reports as well as the supporting research evidence of a commission are sent out for evaluation to all kinds of interested public and private bodies. Thus, when the work of drafting a bill is started in the Ministry, it has at its disposal not only a report, which in most cases is supported by a solid majority within the commission, but also sometimes over a hundred written submissions from the most varied bodies and parties interested in the matter. Particularly in a democracy cherishing consensus this is a procedure that is conducive to it.

Since the commission members who are politicians often have been chosen because they are members of the relevant committee in the Parliament, they can bring the expertise they have acquired after years of work on the commission to bear on the debates about the bill that occur both in the Parliament Committee and when the bill reaches the floor. The Parliament Committee can also call upon experts, for example, educational researchers, to witness. This happened to my knowledge the first time in 1950, when the first Comprehensive Education Bill was under debate and when a 10-year pilot program was contemplated (Husén, 1962). A key issue was the proportion of an age cohort that could profit from academic secondary

education and the actual size of the so-called reserve of ability.

Thus, as far as education is concerned, the liaison between policymakers and researchers has for a couple of decades been a rather close one. The contacts that once were established in the ad hoc commissions have later to some extent become institutionalized by the setting up of various bodies that serve in an advisory capacity within the central agencies in charge of servicing the educational system at various levels. For example, the National Board of Education and its Bureau of Research and Development have an Advisory Council consisting of representatives of various "consumers" groups, such as members of Parliament and teacher associations.

We shall now proceed to a brief description of how lines of communication were opened up and reinforced and how the relations between research workers and policymakers developed after 1940.

The 1940 School Commission approached the holders of the chairs of education and psychology at the Swedish universities and asked for their evaluation of "the present conception of psychological research on mental development in children and adolescents." The commission consisted mainly of school and university teachers, but no members represented educational research, and there were no politicians. Much time was spent on the problem of the proper age or grade level for primary and secondary education, on how to deal with the dualism of the two systems. The commission split into two groups, one in favor of a common 6-year elementary school, upon which all types of further schooling should build, and the other in favor of so-called double transfer, that is, one connection after the fourth and another after the sixth year in the elementary school. The issue boiled down to the problem of the relative "effectiveness" of the four-plus-five and the six-plus-four systems. The commission made extensive evaluations of the two systems in terms of actual number of school years to graduation, number of grade repeaters, social background of students, school marks, and age of graduation (Husén, 1962).

The issue which for decades has been a dominant preoccupation of Swedish as well as other European secondary-school teachers, who were used to teaching an intellectually and socially select group, could schematically be stated like this: At what age would it be necessary for pedagogical reasons to differentiate between the academic and the practical students? With an American vocabulary the problem might be phrased as that of the age at which homogeneous grouping should be introduced or at which the college-bound students should be separated from the rest.

The 1940 Commission wanted these professors to give the evidence of psychological research pertaining to this problem. Except

for the rather limited evaluation inquiries mentioned above, no research was conducted by the commission. The professors relied heavily upon the general literature on developmental research and for the rest fell back upon "the depth of their consciousness."

The appointment of the 1946 School Commission, which consisted mainly of members of Parliament, implied a turning point in various respects. Apart from the fact that the commission politically, and not least in school policy, held rather radical views, it also was the first to launch an extensive research to elucidate the development of "practical" and "academic" aptitudes and their interrelations at various levels.

The commission expressed a strong belief in educational research in its main report which was turned in to the government in 1948: "The findings of educational and psychological research have to be allotted a decisive role as concerns problems which have been subjected to rigid investigations."

The research program was commissioned to Professor John Elmgren at the University of Gothenburg, who called it a study of "ability structure and factorial maturity in Swedish school children." His investigation (Elmgren, 1952) was based on cross-sectional surveys of schoolchildren from 7 to 16 years of age. Elmgren found, like his predecessors, that so-called practical and theoretical aptitudes were positively correlated. Furthermore, according to the factor analyses, "practical" aptitude did not seem to be "as simple and unitary in its structure as general intelligence." The commission drew two conclusions from this in its report in 1948. The earlier conception of theoretical and practical aptitudes as being in a compensatory way related to each other was not valid, and this finding certainly did not support a parallel school structure. An early differentiation by the transfer of the "academics" to a selective secondary school was not justified.

The interesting feature so far of the expectations put by the policymakers upon the researchers was that they should give generalized answers pertaining to school structure and organizational differentiation. They were expected to specify when children's ability structure had reached the level of differentiation that made a separation into different types of schools justified from an educational point of view.

Even though the interaction between school and society was appreciated, this process was thought of in a rather static way. Factors such as social-class differences in the choice of educational and vocational careers were hardly yet realized by the policymakers, and researchers were just beginning to study these problems. Irrespective of

whether those involved had a conservative or a radical outlook concerning school organization, they thought that there was a fixed number of students able to profit from academic instruction at the pre-university and university levels. The dynamic society was still not recognized as an important factor with its insatiable demand for education, nor was education conceived of as an investment in human capital.

The major argument in favor of a reform of secondary schools and higher education was that of providing equality of opportunity. The essence of "democratizing" the schools was to put everyone, irrespective of social, economic, or geographical factors, on equal footing to realize his potential. This initiated two new lines of inquiry—the study of the reserves of ability and research into the social factors that, in a selective educational system, acted as barriers against the realization of the principle of equality of opportunity.

An attempt to come to grips with the "reserve of talent" problem was made by Husén (1948), utilizing data on army conscripts. At that time virtually every young man by the age of 20 had to undergo a psychological and medical examination before he was considered for military service. Since conscripts' records included mental tests, school history and marks, as well as social background, estimates of the talent reserve were made similar to those later made by Wolfle (1954) for the United States. These estimates were quoted by the Minister of Education when he introduced the 1950 Education Bill before Parliament.

Härnqvist (1958), under the auspices of the Royal Commission on Universities, followed one-fourth of the 1945 fourth-grade students up to 1956 or to the age of 20. Based on these estimates of the untapped reserve of talent, strong support was given to more generous policies of student grants and loans for university study.

Investigations of the effects of "social handicaps" were launched in the mid-1940s, some of them projects taken up by the institutions of education and sociology, some sponsored by ad hoc commissions. Boalt (1947) investigated the role of social handicaps in the school careers of a complete age group of youngsters in Stockholm, following them up from the age of 10 to the age of 20. Orring (1959) studied grade-repeating and dropout in the academic secondary school. Husén (1966) reworked the data and related failure to social background. Studies of this type threw doubt on the alleged superiority of a selective and competitive school system in a society with an expanding economy and a strong need for mass education at the upper secondary level. When the relative merits of a selective system and a comprehensive system are weighed against each other, one also has to consider the price that is paid for the quality of the minority

that reaches the top of the ladder. Social bias in the selection and screening procedures in the academic secondary school was obvious.

These investigations played an important role when the 1957 School Commission prepared its report and when the 1962 Education Bill introducing the common 9-year school in all of Sweden was drafted. Similar considerations were also behind the 1964 Act concerning *gymnasium* education.

Shortly after the government had appointed the 1957 School Commission which, according to its terms of reference, had to consider the objectives and thereby the curriculum content of the compulsory school, the Industrial Council for Social and Economic Studies, a private organization supported by the Swedish Employers Association, proposed a research program to be sponsored jointly with the commission, that is, by the government and the Council. According to the initial plans, the curriculum content of mathematics and Swedish was to be the object of the study. But at the request of the commission the scope was widened, and three other subjects—physics, chemistry, and civics—were included. The research program was commissioned to the Institute of Educational Research at the School of Education in Stockholm in partnership with the Council for Social and Economic Studies and the 1957 School Commission. Teachers in the public elementary or secondary schools, as well as teachers in the vocational schools and in the *gymnasia*, were asked to rate topic by topic what requirements they put on the students they received from schools providing basic education. They were also asked to rate the shortcomings. A series of private industrial and commercial enterprises allowed the researchers to give an extensive achievement-test battery to the employees during working time. The employees also completed a questionnaire indicating the frequency of use of the various skills and topics of knowledge under investigation.

An advisory council consisting of representatives of teacher associations, management, and labor unions was set up for the project. The representatives agreed to send letters of recommendation to their members. In order to make appropriate preparations for the field study, which could be expected to cause some concern (among other reasons, because the companies had to pay for a full day's work for thousands of employees who were at the researchers' disposal for 5 hours of achievement testing), a special committee was set up representing the various labor-market organizations. The teacher associations were also represented on the various expert committees that advised on the construction of questionnaires and tests.

The main issue in Swedish school policy from 1940 to 1962 was the problem of school differentiation (Husén, 1962): At what age should

the students considered to be academically talented be separated from their nonintellectual classmates? Should they transfer to another program in a separate school, to another program taught within the same school, or simply to separate classes within the same program in the same school? The theory held by most Swedish secondary-school teachers was that both categories of students would gain from an early separation. The theoretically oriented students would not be hampered by their slow-learning classmates, and the latter would be saved the feeling of inferiority that might be caused by the constant confrontation with their bright peers.

This problem was investigated by the 1957 School Commission, which sponsored an extensive follow-up project reported in a monograph, *Ability Grouping and Scholastic Achievement*, by Svensson (1962). As pointed out above, the study took advantage of a unique situation. The City Council of Stockholm in 1955 decided to embark upon a program with a transitory organization of the 9-year compulsory school in the southern part of the city. Stockholm was thereby divided into two sectors. In the southern sector no separation whatsoever took place during the first 6 school years and in some districts not until the ninth grade. In the northern sector the dual or parallel system was still kept for a short time, which meant transfer for an "élite" to the selective academic school after the fourth grade. Both the northern and southern sections had similar social-class distributions. Since social background and initial ability measured by tests and school marks were kept under control, it was possible to take advantage of the situation and carry out a 5-year follow-up. Comparisons were made between students of all ranges of ability in the two systems. By means of analysis of covariance outcomes could be compared taking initial social and achievement status into consideration.

Marklund (1962) carried out a study on class structure and class size. Since school marks all over Sweden were calibrated by means of the achievement-test batteries in Swedish, mathematics, and English that are given uniformly to all students at certain grade levels, useful survey data were available for various research purposes. Marklund took advantage of these data. Homogeneity (as defined by IQ, school marks, and achievement tests) was studied keeping class size under control and vice versa, an approach that had not been taken by the numerous researchers who had tackled the problem of class size, for instance.

Marklund's study on the school class offers an interesting illustration of how previously evolved lines of communication to policymakers were used and reinforced. His monograph happened to be published a few weeks before the Education Subcommittee of the Budget

Committee in Parliament had to consider a motion by some members of Parliament making provisions for a successive reduction of class size which, by the way, at that time averaged about 25. The rationale spelled out in the motion was that such a reduction would increase educational productivity. Marklund was able to make it quite clear that within the range under consideration there was no evidence to support such a prediction, at least as far as student cognitive competence was concerned.

The role of the National Board of Education ought to be mentioned in this connection. It might be of interest to consider the budget requests that the National Board of Education in 1964 submitted to the government for the program of the Bureau of Research and Development instituted by Parliament in the 1962 Education Act. Parliament every year appropriates funds for the program. The Board then, on the basis of a policy framed as a result of discussions with its Advisory Council, allocates money to research institutions that have been approached by the Bureau or that have turned in proposals. In the budget request there is a lengthy discussion of the suitable percentage of the total educational operational budget to be allocated to educational research; 2 percent was by the mid-1960s suggested as a target for the near future (Svensson & Björklund, 1966). That target has not been achieved. By the mid-1970s it was still below .5 percent.

The Bureau in its budget requests for the fiscal year 1965-66 pointed out that educational research, in order to yield something that could be useful, should be centered around comprehensive programs carried out by research teams that can be kept together for several years. A research director has been assigned to each project, and the junior members of the team are to be given the opportunity of utilizing their work as part of their graduate studies.

With this as background, I see the future of research in education at three different levels and for three different purposes:

First, the universities should pursue their genuine and unique function, which is basic research. This does not imply a withdrawal into an ivory tower, but simply that their researchers should be given sufficient time and resources to develop the new concepts and models without which further research, particularly short-range and expensive developmental work, cannot be productive. It is not the role of the university scholar to develop instructional systems or introduce new classroom practices. His role is properly fulfilled if he stays aloof and indulges in the "irresponsible" thinking that precedes developmental work.

Second, large-scale surveys evaluating new organizational setups introduced by central or local authorities, or projects producing instructional systems which are expected to guide classroom practice, ought to be carried on outside the universities. Since quite a lot of this research deals with the evaluation of new structures or curricula, it is obviously important that the evaluation should be carried out by independent research bodies, such as the National Foundation for Educational Research in England and Wales, or at the Max-Planck-Institut für Bildungsforschung in Berlin. Apart from the fact that, ethically, a governmental or quasi-governmental agency ought not to evaluate its own operations, research within a bureaucratic setting can easily become a nightmare to the researcher. His work is by definition antibureaucratic, informal, innovative, and not limited by timetables.

Third, obviously, certain administrative functions within the central or regional educational agencies cannot be successfully performed unless special research techniques are used to secure data. I am thinking primarily of the planning function, which has to be guided by the collection of information on pupils, teachers, schools, and finance. In some instances, regular sample surveys have to be conducted. In recent years, research in some countries has focused on massive attempts to improve curricula and instructional methods. The tremendous increase of funds that the federal government of the United States is investing in improving school practices is well known to all of us. In countries where a central ministry of education more decisively influences the curriculum—I am not thinking particularly of England—one has begun to realize the possibilities of changing practices by preparing teaching aids or instructional systems. The sheer size of such an enterprise and its financial implications make it a task for a ministry.

We can draw a continuum between the two poles: on the one side, complete aloofness in the ivory tower, and, on the other, total immersion in the noise of action and discussion of issues in the educational marketplace. The problem is that of striking a proper balance—of establishing a good liaison with the policymakers without becoming evangelists and without scaring the other partner into withdrawing his support.

SOME OBSERVATIONS IN RETROSPECT

First, research can help in tidying up problems as they are first encountered by the policymaker. Thirty-three years ago, shortly after I had completed my doctor's degree with a thesis on adolescent problems, I was invited to speak at a parent-teacher association meet-

ing. After the lecture, when questions were invited from the floor, one of the audience asked: "Tell me, Doctor, how late does developmental psychology think that it is appropriate for a fourteen-year-old girl to be out in the evening?" Not too different was a question that the Swedish 1940 School Committee submitted to the professors of education in Sweden: "From the point of view of developmental psychology, at what age would it be advisable to introduce teaching of the first foreign language?" As a young assistant my view did not count, but it was very simple: "Certainly by the age of two, the age when all children in England begin to learn English!"

The problems that are posed are often not only too broadly stated and, therefore, in need of being broken down but, in most cases, they also need thorough reformulation if they are to be backed by research methods. Thus, for instance, in studying the relation between age and language teaching, the cardinal problem is not age but methods of instruction. In 1941, the professors of education replied that the age of 11 would seem most appropriate, thus tacitly assuming that the traditional grammar approach should be used. Since then we have in Sweden started to teach all children English at the age of 10 and even at 9, a decision which some decades ago would have been regarded as pedagogical sacrilege.

The comparisons between the comprehensive and the selective school structure in Sweden during most of the 1950s were naively carried out in terms of the conventional end products. Thus, the "grammar school" students in both structures were compared without consideration of the fact that, at least partly, they pursued different objectives and that they differed considerably in terms of grade-repeaters and dropouts. They were compared, in fact, without reference to differences in the price paid for the end products. Therefore, it seems to me to be of great importance to establish at least occasional contact between researchers and policymakers. The dialogue between them should help to create more articulate conceptions of the problems and their implications. Such a dialogue proved to be an important part of the Stockholm study, mentioned earlier.

Second, more closely at hand and certainly more obvious than the tidying-up task is the "know-how" or technical craftsmanship that researchers are expected to bring to bear on problems that the administrator or policymaker with (or, as a rule, without) the help of the researchers is trying to formulate. The researcher is called upon to do the fact-finding—to carry out the surveys that clarify the descriptive picture of a certain educational phenomenon. The various Royal Commissions in England, such as the Robbins and the Plowden Committees, are cases in point.

One of my co-workers, Sixten Marklund (1962), carried out a study on the relationship between class size and homogeneity and the outcome of instruction. National survey data were used, and, for the first time, size and homogeneity were studied simultaneously: The effect of size was analyzed, keeping homogeneity under control, and vice versa. The monograph that reported the study happened to be submitted a few weeks before the Budget Committee of the Parliament reviewed an amendment to the Education Bill made by some of its members and proposing a successive reduction of class size, which, by the way, in Sweden was and still is one of the lowest in the world. The argument for the motion was that a reduction of size would improve the outcome of instruction, a rationale which was clearly refuted by the study within the class-size interval considered in this particular case, i.e., between 20 and 30 pupils.

Third, the researcher can be called upon to assist in interpreting data that have already been collected. As part of the routine administrative procedures within all ministries of education, hosts of statistical information are collected. These data are particularly useful in improving planning within the ministry. The researcher can assist not only in interpreting data already gathered but also in advising what might be collected, so that more relevant aspects of the educational reality are covered by the statistics.

In recent years the planning function has grown rapidly to such an extent that research has become part and parcel of planning. I am not thinking primarily of fancy mathematical models which seem to shun encounters with empirical reality, but rather of the basic surveys that are needed.

In my own country as well as in other countries I have encountered some confusion about the proper role of the researcher in the decision-making process. For instance, it has been argued that decisions about the school structure should not be made on the basis of "pure political deliberations." When in 1962 the Swedish Parliament passed legislation that made the introduction of the 9-year, undifferentiated comprehensive school mandatory all over the country, the objection was raised that this was a "political" and not an "expert" decision. A considerable number of the secondary-school teachers who had been trained for an elitist system resented the fact that their "expert" opinion was overruled.

It has often been stated that wars are too important to be run by generals. It has less often been emphasized that schools are too important to be run by educators, whether teachers, school administrators, or professors of education. The appropriate role of the educator in the decision-making process is primarily to provide information and

expert judgments and thereby to help broaden the base upon which decisions are made by those who, in a representative democratic system, have been elected to make these decisions. If the educator is a researcher, it is even more important that he should play his role according to the rules and confine it to professional performance as long as he wears his researcher's hat. Let me hasten to add that the educator, whether a teacher or a researcher, is obviously entitled as a private person to advance opinions or make political statements. Equally self-evidently, however, he should try to keep these personal statements clearly separate from the generalizations drawn from his scholarly inquiries.

The nature of the research can easily affect the relationship between the researcher and the policymaker. Problems pertaining to school structure are, since they have profound implications for the educational opportunities of groups of individuals, issues that easily can cause political temperatures to rise and governments to be shaken. This brings considerable pressure to bear upon the researcher. The directive given to him might be confined to the study of certain aspects of a problem, depending upon the terms of reference of the research. The temptation to spread an approved "gospel," whether he deeply believes in it or not, is particularly strong if the researcher (or the institution to which he belongs) is entirely dependent on the financial support of the agency that commissioned the project. Instead of confining his role to research, and studying the problem with the aim of increasing understanding, he easily becomes an evangelist and not a healthy "skeptic" to whom the slogan of the day or the policy statement is a hypothesis, which, at best, lends itself to scientific inquiry.

REFERENCES

- Boalt, G. *Skolutbildning och skolresultat för barn ur olika samhällsgrupper i Stockholm*. Stockholm: Norstedts, 1947.
- Elmgren, J. *School and psychology*. Stockholm: Statens Offentliga Utredningar, 1952.
- Härnqvist, K. *Reserverna för högre utbildning* (Rep. No. 11). Stockholm: Statens Offentliga Utredningar, 1958.
- Husén, T. *Begåvning och miljö*. Stockholm: Almqvist & Wiksell, 1948.
- Husén, T. *Problems of differentiation in Swedish compulsory schooling*. Stockholm: Scandinavian University Books, 1962.
- Husén, T. The relation between selectivity and social class in secondary education. *International Journal of Educational Sciences*, 1966, 1, 17-27.
- Marklund, S. *Skolklassens storlek och struktur*. Stockholm: Almqvist & Wiksell, 1962.

- Orring, J. *Flyttning, kvarsittning och utkuggning i högre skolor i relation till folkskolans betygsättning* (Rep. No. 35). Stockholm: Statens Offentliga Utredningar, 1959.
- Svensson, N. E. *Ability grouping and scholastic achievement*. Stockholm: Almqvist & Wiksell, 1962.
- Svensson, N. E., & Björklund, E. Educational research and development in Sweden. *Educational Sociology*, 1966, **39**, 145-161.
- Wolfe, D. *America's resources of specialized talent*. New York: Harper, 1954.

Chapter 9 — Summary

EDUCATION, SOCIAL SCIENCE, AND THE LAW: THE RODRIGUEZ CASE AND SCHOOL FINANCE REFORM

MARVIN BRESSLER

Princeton University

and

JOEL S. BERKE

Educational Policy Research Institute

Educational Testing Service

DISPUTES over how best to generate and allocate funds for tax-supported schools are as old as public education. They involve the interplay of philosophical concepts like equality and quality in education with the most direct of vested interests in tax bills and the availability of educational services. Since the late 1960s that debate has been conducted, at least in part, in the courts. Since judicial verdicts rest on both fact and legal principle, the protagonists often enlist the aid of social scientists, whose theories and research provide arguments for all parties to the litigation.

One aspect of this recent "judicialization" of educational policy conflict is the highlighting of a fascinating and significant triangular relationship between education, law, and social science. No court proceeding is more instructive in this regard than *San Antonio v. Rodriguez* (411 U.S. 1, 1973), a case of immense social importance which could have required the transformation of arrangements for raising and distributing educational revenues in every state except Hawaii. This chapter describes those interrelationships and attempts

to trace their implications for broader issues of social science and policy development in school finance.¹

After a synopsis of *Rodriguez*, we review the main traditions of scholarship in school finance. The occasionally stormy marriage of law and social science in trying to win school-finance reform accompanies the controversy over seven major issues in the Texas case. The judicial reasoning it reflected is observed in the courts' chosen sources of authority and criteria of relevance.

In legal proceedings, attorneys and social scientists alike perform as partisan advocates, but this role creates more dilemmas for the social scientist than for the legal professional. We attempt to describe and suggest modes of dealing with these dilemmas, and close by reviewing recent developments in social advocacy, research, and public policy that affect the prospects of school-finance reform.

Rodriguez began when parents from the poorest of the 12 independent school districts in San Antonio sued state and county officials, charging that their district suffered from a statewide pattern wherein the state subsidized the rich at the poor's expense. Each district received about the same sum per pupil in state aid, but the property-rich districts were able to raise many hundreds of additional dollars per pupil from local tax rates nearly nine times lower than rates that, in poor districts, raised only a few additional dollars. Social scientists marshalled data to show these interdistrict inequities not only through aggregate dollars but also through variables such as levels of teacher training and counselor-student ratios.

In 1971 a federal district court ruled unanimously for the parent group, striking down Texas' school-finance laws as a denial of the Fourteenth Amendment. The state legislature was told to frame laws that would no longer base the quality of a child's education upon the wealth of his or her school district. But Texas appealed the decision, and in March 1973 the U.S. Supreme Court narrowly overturned the earlier decision, by five to four. The majority—Justice Powell, joined by Justices Blackmun, Burger, Rehnquist, and Stewart—acknowledged that the Texas system needed reform but denied that it violated equal-protection standards of constitutionality. Despite forceful dissents from Justices Marshall, Douglas, Brennan, and White, this decision doomed the possibility of using the Fourteenth Amendment's equal protection standard in federal litigation to eliminate fiscal manifestations of unequal educational opportunity.

Suits pending in more than 20 states indirectly rested on this

¹ The chapter itself contains two appendices that reproduce the evidentiary affidavits that include the factual analysis made by the judges.

argument, and the lower court's opinion in *Rodriguez* had been nearly identical to decisions in California, Minnesota, and three other states which consequently began to alter their finance systems. After the defeat in *Rodriguez*, litigants for reform now perforce trained their sights on state courts and constitutions.

From the outset, advocacy and eclecticism have characterized school-finance scholarship. Although by the 1950s the rural and small-town bias of educational formulas drawn from early scholarship was apparent, the conceptual base of this scholarship continues to influence reformist plans. The eclectic quality of scholarship in the field, beginning in educational administration, adapting economics to its purposes and most recently incorporating legal scholarship, has sought to produce a better harvest of legislative solutions.

Elwood Cubberley's turn-of-the-century studies of the property tax base in several states begin the tradition and show its normative concern with equity in taxation as well as equality in provision of services. The *foundation plan* conceived by Strayer and Haig in 1923 used state grants-in-aid to equalize disparities between districts in taxable resources in order to finance a state-defined minimum school program in each district. The *percentage-equalizing grant* of the same period allowed localities to select their own levels of school support, while the state shared costs in inverse proportion to the wealth of the local property tax base. The other contribution of the early reform scholarship to exert continuing influence is Henry Morrison's *full state assumption* of educational costs and revenue procurement. Proposed in 1930 but adopted only by Hawaii in the late 1960s, the concept perennially appeals to state commissions and special studies of school-finance reforms.

The eclecticism of school-finance scholarship increased in the 1960s, when economists led by Charles Benson and Jesse Burkhead were drawn to the topic as an important public function in the larger context of American political economy. Focusing on empirical and theoretical inquiry into systems of taxation and intergovernmental flows of funds, economists contributed five important approaches to the study of financing public education.

The work of Richard Netzer focused on the economics of the property tax, particularly early issues of inequitable administration and regressivity in incidence. Such work has continued to the present, with both policy applications and lively reinterpretations.

Nationwide demographic and economic shifts since World War II and their damaging effects upon the nation's largest cities, which were left with inadequate tax bases and a more than adequate share of needy consumers of social services, were documented in studies by

Harvey Brazer, Alan K. Campbell and Seymour Sacks, and others. Such studies have subsequently prompted the inclusion of "urban factors" and income measures in school-finance formulas.

Distributive issues engaged economists like Werner Hirsch. Does intergovernmental aid increase expenditures by recipient governments or distort resource allocation in other ways? This line of inquiry promoted examinations of types of state aid to local school districts and used the economic principles of income and substitution effects as guides to further inquiry.

A fourth approach useful to school-finance studies seeks the determinants of public expenditure decisions. The findings of this work by scholars such as James, Kelly, Garms, and others establish a strong relation between property valuation, socioeconomic characteristics, and district expenditure levels for education and other functions. Analyses of this type have been instrumental in demonstrating the unconstitutionality of numerous state aid systems.

Fifth, Burkhead, Thomas, and others applied to public education some analytic methods used to examine relations between factors of production and industrial outputs. But the recent research of Wiley and Harnischfeger suggests that when process factors and time of exposure capture the real experiences of particular students, effects of schooling become significant.

About 1965, legal scholarship began to feed the eclectic stream of school-finance studies. It was urgent to close the distance between research findings like those of the economists and the legislation that would be needed to transform existing school-finance systems.

Arthur Wise in *Rich Schools, Poor Schools* (1968) suggested that major constitutional precedents on the Fourteenth Amendment's equal-protection clause applied to existing state school-finance systems would result in a finding that such laws were unconstitutional. Wise's theory was first tested in *McInnis v. Shapiro*, an Illinois suit that was dismissed by the federal district court for failure to state appropriate constitutional grounds for relief. The court admitted, however, that the inequalities of Illinois arrangements were "readily apparent." The court was troubled by its judgment that the allocation of public money and the definition of "educational need" are policy decisions better left to legislatures, not to courts.

John C. Coons' contribution exemplified another strategic approach to school-finance reform through court action. He and his colleagues William Clune, III, and Stephen Sugarman in *Private Wealth and Public Education* (1970) suggested the litigation strategy that would win in *Serrano* and most other cases. Their theory, known as the test of fiscal neutrality, held that only the wealth of a whole

state, not merely of its individual school districts, can justly condition the sum available for a child's education.

The "power-equalizing" approach of Coons aimed primarily at achieving tax rate equity so that any level of tax effort would produce equal education money for all school districts at that rate. He and his colleagues emphasized the principle of "subsidiarity," the right of individual units to select different levels of educational services freely, unimpeded by differences in wealth. Philosophically, Wise's approach was more egalitarian. It urged at least equal treatment of all students and, at best, a compensatory approach that would direct public resources toward overcoming student disadvantages born of social and economic deprivation. Both approaches became important in the *Rodriguez* litigation.

Although the outcome of *Rodriguez* depended upon the applicability of the equal-protection clause to the method of financing public education, no one challenged the established legal framework of equal-protection argumentation. It is worth pausing briefly to note key terms of equal-protection logic before assessing the seven major issues in the Texas case. Justice Powell summarized the analytic sequence necessary:

We must decide, first, whether the Texas system of financing public education operates to the disadvantage of some suspect class or impinges upon a fundamental right explicitly or implicitly protected by the Constitution, thereby requiring strict judicial scrutiny. If so, the judgment of the District Court should be affirmed. If not, the Texas scheme must still be examined to determine whether it rationally furthers some legitimate, articulated state purpose and therefore does not constitute an invidious discrimination in violation of the Equal Protection Clause of the Fourteenth Amendment. (*San Antonio v. Rodriguez*, 411 U.S. 1, 17 [1973])

We will take the key terms in order:

Suspect class. Ordinarily, the law is guided by the doctrine of "presumptive constitutionality" in judging the actions of legislatures, even in the numerous instances when a legislature discriminates for legal purposes between groups of persons (as in denying permission to drive to persons under a certain age). Judicial precedent disallows some kinds of discriminations, regarding them as "suspect" and inviting "strict scrutiny," for example, in discriminations based on race, nationality, religion, wealth, and sex. Under scrutiny, a state may be required to show that it has not violated a "fundamental right" of the designated population.

Fundamental right. This concept is not synonymous with social importance or distributive justice. More is required to be shown than that various forms of deprivation suffered by a group injure its life

chances, prestige, or power. To qualify as fundamental, a right must be explicitly or implicitly guaranteed by the Constitution. Unequal access to housing may be a personal misfortune, but since it is judged not to be constitutionally guaranteed, relief must be sought politically, not judicially. Yet a poll tax, arguably no more harmful to individual welfare than closed access to housing, violates the First Amendment and is judged subject to the courts.

Compelling state interest. If a state can establish that a discrimination it legislates is based on paramount values that make even stronger claims than the doctrine of equal protection of the laws, it may continue to reward some groups and penalize others. Hence Justice Powell's statement of the need to determine whether the Texas system furthers some "legitimate, articulated state purpose." The term "compelling," however, implies that the total collectivity's interest may supersede the rights of constituent groups only under circumstances very rare indeed. The questionable imprisonment of Japanese-Americans during World War II is probably the most dramatic example of a violation in the name of "compelling state interest."

In applying equal-protection analysis to *Rodriguez*, the Supreme Court identified seven major educational issues bearing on the case. We will take the questions one at a time, noting the interplay of law and social science in the court proceedings.

1. *May a state distribute its educational resources unequally on the basis of the unequal wealth of its recipients?*

The main empirical evidence social scientists developed for *Rodriguez* was intended to demonstrate that Texas' method of raising and distributing educational money created a "suspect class" of poor children from poor districts, discriminated against by making the quality of education a function of parental and neighborhood wealth. To the three judges of the district court the evidence was unanimously persuasive: "For poor school districts educational financing in Texas is, thus, a tax more, spend less system," they wrote (*Rodriguez v. San Antonio*, 337 F. Supp. 280, 282-283 [WD Tex. 1971]).

Based on a stratified random sample of 110 of Texas' 1,183 districts, the financial study directed by Joel S. Berke, then of Syracuse University Research Corporation, revealed a linear relationship between the amount of a district's taxable property and its per-pupil expenditure for education. The Supreme Court majority noted a critique of the methodology used in the study, but accepted its basic conclusions "insofar as any . . . is relevant to the Constitutional thesis." As we shall see, relevance becomes the nub of the issue.

The fiscal bias data were central to the dissenting opinion; the doctrine of "fiscal neutrality" was attacked by the attorney for Texas

as offensive to the essence of democratic capitalism and tending toward "socialized education." But even the majority of the Court was prepared to acknowledge that wealth-based discrimination was constitutionally suspect. The dispute focused on whether interdistrict disparities in per-pupil expenditure, in and of themselves, violated Fourteenth Amendment guarantees.

No legal problem so divided the court as that of defining the class that required constitutional protection. The majority insisted that existing precedents establishing wealth as suspect applied to poor individuals, not to poor districts. The issue was not the data, with which all parties were content, but the doctrine.

2. *Shall poor children or children living in tax-poor school districts be defined as the class requiring constitutional protection?*

Both majority and minority of the Supreme Court agreed that substantial correlation existed between district property wealth and per-pupil expenditures. To Justice Marshall, that relation testified to the existence of unconstitutional discrimination; to Justice Powell it did not. For Powell to find the Texas system unconstitutional it would have been necessary to show an inverse relationship between *personal* income and school expenditures throughout the range of districts in the sample. The selection of relevant information, not the validity of the statistical relationship, determined the issue.

3. *Is education "fundamental" or merely important to the welfare of the nation and its citizens?*

Equal schooling is protected by the Fourteenth Amendment only if linked to rights that are guaranteed elsewhere in the Constitution. Since nowhere in the Constitution is education directly mentioned, the plaintiffs were obliged to claim (and Texas to deny) that equal educational opportunity is a prerequisite to the exercise of other, indisputably fundamental rights like free speech. It was argued against the state that, while the First Amendment "protects uninformed, unpersuasive, and possibly unintelligible speech, the meaningful exercise of that right . . . is dependent upon [the speaker's] ability to speak intelligently and knowledgeably" (pp. 31-32). The plaintiffs cited social science data that strongly correlate educational attainment with political participation, especially voting behavior. The minority opinion shared this broad construction of the First Amendment, conceiving free speech as a meaningful, not merely a formal, right. By contrast, the Powell opinion took a more orthodox, less substantive view of the free speech guarantee and its link to education: "We have never presumed to possess either the ability or the authority to guarantee to the citizenry the most *effective* speech or the

most informed electoral choice. . . . These are indeed goals to be pursued. . . . But they are not values to be implemented by judicial intrusion into otherwise legitimate state activities" (411 U.S. 1, 36).

Again, on this question, the data were not denied; the Justices simply chose to bypass the factual issues. In citing social science evidence, the minority position risked rejection because of demonstrably faulty or ambiguous findings; in a kind of double jeopardy, it also risked rejection of findings as irrelevant.

4. *Shall educational quality be defined by the resources made available to children or by the outcomes that are demonstrably attributable to schooling?*

In translating the economists' "input-output" question into the language of education, the data seemed to show that, when individual and social characteristics of pupils were held constant, the school system had relatively little effect on measures of cognitive achievement. At every stage of the case, Texas relied on the research of James Coleman and others that doubted the link between resources and student achievement. (Since these observations were made of existing schools, of course, the data could as easily have supported a case for massive spending on school improvement.) Seizing the popular implication that "money does not matter," Texas argued that pupils in the poorer district could not therefore be denied the hollow benefits of "quality education" as measured by resources.

Now it was the minority's turn to claim that even if the Coleman data were valid, they were irrelevant to the constitutional issues. Other data were brought forward to demonstrate "a positive relationship between educational funding and learning, whether the latter is measured solely in terms of cognitive achievement or in terms of other qualities such as creativity and socialization" (p. 13).

In short, the evidence was employed by both majority and minority to add force and flourish to arguments which, in their opinion, could stand alone even if the social science findings were reversed. Justice Marshall's minority opinion does not address the validity of either version of the cost-quality controversy; he makes the legal point that the burden of proof that cost disparities do not affect the quality of education falls on the State of Texas, which nonetheless never tried to show that educational quality is unaffected by scanty resources. Similarly, Justice Powell found the cost-quality argument peripheral. He could have accepted that cost determined quality without retreating from the majority view that the Texas system was constitutional.

5. *Shall the quality of education, however defined, be conceived as equal where it is tailored to the child's individual needs, is the same as education offered to the most privileged, or satisfies the requirements of a specific adequate standard?*

Both sides were obliged to define the meaning of "equality" in education, once each had identified its chosen measure of quality in schooling—for the majority, educational outcomes in "basic minimal skills" (411 U.S. 1, 37); for the minority, educational investments.

Since all Texas schools presumably instructed pupils in basic skills, the majority position in effect defined equality in expenditures as merely matching the spending level of the state's poorest districts. The minority defined equality as comparable real expenditures for all children and all districts.

On this question, as on others, the court essentially resolved the issue on its judicial and philosophical merits alone. Whatever social scientist advocates believed, the facts of the existing situation in Texas were irrelevant.

6. *What are the positive and negative consequences of a school-finance system which relies on the local property tax for the general social and educational goals of the state?*

Texas argued that, in effect, a locally based financing system has wholly positive consequences. Keeping the local property tax as the chief fiscal instrument for education is vital to the survival of local control over educational policy. Local control became the majority's "compelling state interest." He who pays the piper calls the tune, they reasoned, and concluded that fiscal neutrality would weaken traditional local prerogatives such as choosing curricula and allocating funds.

One *amicus* brief which impressed the Court argued that the lower court's decision unjustly limited the right of a parent "to pay more for his child's education" (Brief *Amicus Curiae* of Richard M. Clowes, Superintendent of Schools of the County of Los Angeles, et al., pp. 35-36). Mr. Justice White's dissenting opinion countered that the existing system "utterly fails to extend a realistic choice to parents" in the poorer school districts. For these parents, already taxed to the hilt for inadequate school services, the property tax is legally and practically no longer available as a source of funds. In this situation, "local control," in the sense of a large menu of educational program choices, is merely fanciful.

The minority weakened its case severely by not challenging the sanctity of local control. Legally, it could have been asserted that in public education, fiscal neutrality and equal educational opportunity are compelling state interests that outweigh all others.

Better yet, social science evidence could have shown that in reality, centralized financing and localized policymaking do not inherently conflict. Detailed accounts of the Hawaiian system might have demonstrated this, though it is not difficult to imagine the jousting about

sample adequacy, flaws of design, and generalizability that would have followed. Still, it is curious that social science was not injected into the case at the point where it might have exerted maximum strength, arraying fact to counter a traditional fiction of American federal politics.

7. *Are there alternative methods of financing schools that could remedy the defects of present systems while retaining their desirable features?*

Justice Powell's opinion admits the need for reform in state systems of financing education but reveals the less interventionist character of the Supreme Court by underscoring its reluctance to dictate public policy, the proper province of legislatures: "The ultimate solutions must come from the lawmakers and from the democratic pressure of those who elect them" (411 U.S. 58, 59). Justice Marshall, on the other hand, was "unsatisfied with the hope of an ultimate 'political' solution" while time and continuing inferior education erode the prospects of a generation of youngsters. Marshall suggested a number of plans drawn from the school-finance literature, but the case for any was unconvincing to the majority, which was quick to point to the disagreements over their effectiveness and constitutionality among reformist scholars.

It would be easy but incorrect to deduce from this history that social research has had negligible effect on school-finance litigation. The remainder of this chapter attempts to suggest lessons from the dilemmas and selectivity among findings that *Rodriguez* revealed.

First, social scientists who would play advocacy roles should understand how the judiciary ordinarily establishes the existence of a fact. In *Rodriguez* there were three sources of evidence: common knowledge, judicial authority, and social research. The Court was most inclined in this case to rely on judicial authority, secondarily on propositions based in conventional wisdom. This is hardly astonishing, for judges must rule and no calendar would ever be cleared if courts tried to become research institutes for each proceeding. But the social scientist in court must ask, first, under what circumstance any fact will be important in resolving a legal issue and, second, to what extent the facts adduced by social science will be clearly superior to other sorts of knowledge.

The criterion of relevance, as we saw, pushed aside many social science findings in *Rodriguez*. Had correlations between school expenditures and socioeconomic characteristics of Texas' population been perfect or even unusually high, the Court might have affirmed that wealth was a "suspect category." But social science findings are seldom airtight, and evidence can be used selectively to deny the relevance

even of undisputed fact. In short, the data can often establish plausibility for empirical assertions, but these may not be regarded as germane for resolving legal issues.

Even when a class of fact is admittedly relevant, social science evidence may be discounted by the court because of the state of the art. A social-scientific proposition introduced in a legal proceeding should be clear, certain, stable, and framed at the appropriate level of abstraction—an ideal seldom realized. Failures of communication are legion, given the arcane technical vocabularies of many social sciences. Not unjustly, conscientious but puzzled jurists must then fall back on "simple arithmetic and morality," as Judge Skelly Wright put it, in rendering a decision. In *Hobson v. Hansen*, Wright added: "The lawyers in this case had a basic responsibility which they have not completely met to put the hard core statistical demonstrations into language which serious and concerned laymen could, with effort, understand" (327 F. Supp. D.D.C., 1971 at 859).

Social science is most successful at the descriptive level and when it employs simple variables, as in showing an inverse association in San Antonio between district school expenditures and local tax effort, one that was especially strong at extreme ends of the income distribution. But social science is less able to enlighten complex concepts such as "equality" and "quality," where operational definitions are elusive and measurement problems great. It is almost unable to generate firm general theories at a high level of abstraction. Small wonder that the Court, accustomed to deriving large truths from general legal principles, withdrew to the familiar comfort of judicial authority and common knowledge.

Opinions of social scientists that are otherwise uncorroborated by well-confirmed theory or research should have the same legal status as the reflections of other thoughtful people, not least because the values of any scholar inevitably intrude on his or her work. The academic community, overwhelmingly "liberal" and "enlightened," collectively embraces beliefs that respond just as much to a shifting Zeitgeist as do the clichés of other citizens. And social scientists, eager to solve social problems, introduce evidence in some cases as if it could resolve disputes over moral or philosophical problems.

The current infirmities of all social disciplines, the ambiguous nature of their findings, their vulnerability to faddism, and their capacity for transmuting moral issues into empirical debates should give ample grounds for questioning the role of social science as an agent of educational reform through the courts.

We believe that the social scientist can most honestly deal with the dilemmas raised by participating in legal proceedings by acknowledg-

ing his or her forensic role as a partisan advocate.² Adopting Alice Rivlin's concept of a "forensic social science" which abandons any claim to be part of the "tradition of balanced, objective social science in which the scholar hides (or claims to hide) his personal biases, and attempts to present all the evidence," scholars would write briefs which concede little to the opposition.

This conception may seem antagonistic to any known code of scholarly ethics. Yet intellectual honesty derives its meaning from specific contexts; it is a relation between a scholar's behavior and his *professed* standards. Thus what is inappropriate in a classroom may be entirely proper in a courtroom.

What can be done to clarify the separate roles of social scientist as scholar and as advocate? Procedural mechanisms, including panels of consultants selected from lists submitted by professional associations, might help judges to evaluate conflicting scholarly testimony. The verdict of a judge thus protected might reflect more impartial norms of social science—or at least convey a consensus about reputability—but the jurist would have no cause to rebuke social scientist advocates for speaking partial truths.

Arthur E. Wise identified four principal areas in which social science advocacy can aid the judicial resolution of issues: fact determination, predicting the consequences of decisions, formulating remedies, and evaluating implementations. Since social advocacy is energized by preferences for particular policies, it can also fashion comprehensive strategies to seek its goals.

School-finance reform's prospects have been materially affected by recent developments in litigation, research, and legislation—the three-pronged effort we have labeled the school-finance reform movement. In closing we review them.

Though legal challenges to existing systems failed at the federal level, state courts in at least eight instances have sustained them. The indirect impact of these decisions on public awareness and political action is perhaps most important. The 18 months between *Serrano* I and the Supreme Court decision in *Rodriguez* prompted a host of state commission and legislative studies that brought new laws. If the reform wave has abated somewhat since 1975, increasingly austere state budgets were probably most responsible.

² A colloquy among lawyer-scholars prominent in school-finance litigation after the promising *Serrano* decision is remarkable for its lack of pretense that participants were seeking "objective" truth. They were discussing how to use social science findings most effectively in making the reformist case. Reprinted in the *Yale Review of Law and Social Action*, Vol. 2, No. 2 (Winter 1971), pp. 112–113, we quote it extensively in the chapter of which this is a summary.

The post-*Rodriguez* suits were targeted at state constitutions. Only weeks after the Supreme Court decision, the New Jersey Supreme Court in *Robinson v. Cahill* found that the state funding system violated state constitutional guarantees. In winter 1975, a Connecticut court ruled that the state's funding system violated both the equal protection and the free public education provisions of the state constitution; the decision was sustained in 1977 (*Horton v. Meskill*). In what may have been the final round of *Serrano*, the California Supreme Court in 1977 reaffirmed its earlier decision, this time in terms of the state constitution which, like the constitutions of most other states, explicitly establishes education as a fundamental right.

Grounds for the three rulings differed significantly. California and Connecticut based their findings on the state equal-protection clause, the fundamental nature of education, and "fiscal neutrality," i.e., the impropriety of resting a fundamental right upon the ephemeral size of the local tax base. In contrast, New Jersey rejected the equal-protection, fundamental-right position, but found that the state constitution guaranteed a "thorough and efficient" educational system, thereby requiring the state to provide equal educational opportunity. The fact that all state constitutions have equal protection and education provisions of some kind suggests that school-finance litigation will not lose all momentum.

Research continues its prominent role in this protracted reform drama. Legislatures and national education and legislative lobbies maintain their interest in keeping abreast of new findings. Increasingly, the unifying commitment of school-finance scholars drawn from economics, finance, and the law is devoted to policy analysis. Five types of studies have shaped new policies for school finance.

Analyses of fiscal needs and resources have shown that intercommunity disparities in taxable capacity, tax burdens, and yields offer a rationale for new laws. Substate regional analyses and studies of varying pupil populations, such as the handicapped, have suggested other types of special tax treatment or special distributional provisions.

Second, policy studies of aid formulas illustrate the failure of traditional school-finance law to meet these problems. Formulas devised in an era of wealthy central cities now disproportionately aid already wealthy suburban districts. Further, nominally equalizing formulas often have so many limits put upon their operation that the pattern of spending is highly correlated with property valuation per pupil.

Third, policy analysts have given substance to different value approaches by devising the technical mechanisms for realizing them.

States that want to retain local choice of tax rates can legislate "equal yield for equal tax effort," while states that hold equity in matching resources to needs as paramount can legislate other mechanisms.

Fourth, computer-simulation technology is now used to project the likely effects of potential formulas. In consequence, policymakers can draft legislation with considerably more assurance about its tax and distributive effects. By merging school-finance data with information on general municipal funding and demographics, scholars can anticipate the impact of a school-finance reform upon many other aspects of public policy.

Finally, research on the political systems through which legislative reform will be enacted and enforced can backstop the necessary political effort to assemble coalitions to win legislation that is substantively sound as well as politically palatable. The cumulative effect of reform experience in many states is helpful; for example, we know the advantage of the legislative process over referenda in winning reform, the link of study commissions to political dominance, and the preeminence of political (not educational) leadership and interest groups in determining the fate of reform.

Recent legislation reflects the same competing approaches visible in the traditional formulas and in the litigation we have examined. On the tax side, will a state provide equal yield for equal effort among districts? Will the state mandate uniform property tax rates for education or use statewide revenue sources? On the expenditure side, the new laws attempt to achieve nearly uniform per-pupil expenditure in every district, to vary expenditure levels with presumed differences in service needs, or (in states with "equal yield for equal effort" formulas) to ensure expenditure levels that reflect fiscally neutral local preference.

Even federal aid has been touched by these developments. Several categorical formulas have been adjusted toward congruence with school-finance reform principles, and planning money to help develop state school-finance laws was made available.

In the 1970s, significant changes were made by between 12 and 20 states (depending on how rigorously one defines "significant") in their modes of financing education. These perceptible signs of progress against a backdrop of the old patterns must encourage us. They are due in no small measure to a generation of social advocates for whom setbacks such as the *Rodriguez* case are skirmishes in a long campaign.

REFERENCES

- Coons, J. E., Clune, W. H., III, & Sugarman, S. D. *Private wealth and public education*. Cambridge, Mass.: Harvard University, Belknap Press, 1970.
Wise, A. E. *Rich schools, poor schools: The promise of equal educational opportunity*. Chicago: University of Chicago Press, 1968.